

PCT/CA/ 99 / 01147

1 MAR 2000 (01.03.00)

PA 209398

REC'D 14 MAR 2000

WIPO

PCT

# THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

February 15, 2000

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE UNDER 35 USC 111.

APPLICATION NUMBER: 60/110,427

FILING DATE: December 01, 1998

## PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)



By Authority of the  
COMMISSIONER OF PATENTS AND TRADEMARKS

P. SWAIN

Certifying Officer

Please type a plus sign (+) inside this box → ☐

PTO/SB/16 (12-97)  
Approved for use through 1/31/98. OMB 0651-0037  
Patent and Trademark Office, U.S. DEPARTMENT OF COMMERCE  
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

**PROVISIONAL APPLICATION FOR PATENT COVER SHEET**  
This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53 (c).

INVENTOR(S)					
Given Name (first and middle (if any))		Family Name or Surname		Residence (City and either State or Foreign Country)	
Andrew D		Murdin		146 Rhodes Circle Newmarket	
Raymond P		Oomen		Ontario, Canada L3X 1V2 RR No 1 Schomberg, Ontario, Canada	
<input type="checkbox"/> Additional inventors are being named on the _____ separately numbered sheets attached hereto					
TITLE OF THE INVENTION (280 characters max)					
Chlamydia Antigens and Corresponding DNA Fragments and Uses Thereof					
Direct all correspondence to:			CORRESPONDENCE ADDRESS		
<input type="checkbox"/> Customer Number			Place Customer Number Bar Code Label here		
OR			Type Customer Number here		
<input checked="" type="checkbox"/> Firm or Individual Name			Gavin R. Zealey		
Address			1755 Steeles Ave West		
Address					
City		State	Ontario	ZIP	M2R 3T4
Country		Canada	Telephone	416-667-2854	Fax 416-667-2860
ENCLOSED APPLICATION PARTS (check all that apply)					
<input checked="" type="checkbox"/> Specification Number of Pages		31	<input type="checkbox"/> Small Entity Statement		
<input checked="" type="checkbox"/> Drawing(s) Number of Sheets		13	<input type="checkbox"/> Other (specify) _____		
METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT (check one)					
<input type="checkbox"/> A check or money order is enclosed to cover the filing fees				FILING FEE AMOUNT (\$)	
<input checked="" type="checkbox"/> The Commissioner is hereby authorized to charge filing fees or credit any overpayment to Deposit Account Number: 50-0244				150.00	
The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.					
<input checked="" type="checkbox"/> No.					
<input type="checkbox"/> Yes, the name of the U.S. Government agency and the Government contract number are: _____					

Respectfully submitted,

SIGNATURE

Date 10/27/98

TYPED or PRINTED NAME Gavin R. Zealey

REGISTRATION NO. 39,475

TELEPHONE 416-667-2854

(if appropriate)  
Docket Number:

RY- 49

**USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT**

Burden Hour Statement: This form is estimated to take 0.2 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, Patent and Trademark Office, Washington, DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Box Provisional Application, Assistant Commissioner for Patents, Washington, DC 20231.

TITLE OF INVENTIONCHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND USES  
THEREOFFIELD OF INVENTION

5           The present invention relates to *Chlamydia* antigens and corresponding DNA molecules, which can be used in methods to prevent and treat *Chlamydia* infection in mammals, such as humans.

BACKGROUND OF THE INVENTION

10           Chlamydiae are prokaryotes. They exhibit morphologic and structural similarities to gram-negative bacteria including a trilaminar outer membrane, which contains lipopolysaccharide and several membrane proteins that are structurally and functionally analogous to proteins found in *E coli*. They are obligate intra-cellular parasites with a unique biphasic life cycle consisting of a metabolically inactive but infectious extracellular stage and a replicating but non-infectious intracellular stage. The replicative stage of the life-cycle takes place within a membrane-bound inclusion which sequesters the bacteria away from the cytoplasm of the infected host cell.

15           *C. pneumoniae* is a common human pathogen, originally described as the TWAR strain of *Chlamydia psittaci* but subsequently recognised to be a new species. *C. pneumoniae* is antigenically, genetically and morphologically distinct from other chlamydia species (*C. trachomatis*, *C. pecorum* and *C. psittaci*). It shows 10% or less DNA sequence homology with either of *C. trachomatis* or *C. psittaci* and so far appears to consist of only a single strain, TWAR.

20           *C. pneumoniae* is a common cause of community acquired pneumonia, only less frequent than *Streptococcus pneumoniae* and *Mycoplasma pneumoniae* (Ref 1,2). It can also cause upper respiratory tract symptoms and disease, including bronchitis and sinusitis (Ref 1,3,4,5). The great majority of the adult population (over 60%) has antibodies to *C. pneumoniae* (Ref 5), indicating past infection which was unrecognized or asymptomatic.

25           Of considerable importance is the association of atherosclerosis and *C. pneumoniae* infection. There are several epidemiological studies showing a correlation of previous infections with *C. pneumoniae* and heart attacks, coronary artery and carotid artery disease (Ref 6-10). Moreover, the organisms has been detected in atheromas and fatty streaks of the coronary, carotid, peripheral arteries and aorta (Ref 11-15). Viable *C. pneumoniae* has been recovered

from the coronary and carotid artery (Ref 16,17). Furthermore, it has been shown that *C. pneumoniae* can induce changes of atherosclerosis in a rabbit model (Ref 18). Taken together, these results indicate that it is highly probable that *C. pneumoniae* can cause atherosclerosis in humans, though the epidemiological importance of chlamydial atherosclerosis remains to be demonstrated.

A number of recent studies have also indicated an association between *C. pneumoniae* infection and asthma. Infection has been linked to wheezing, asthmatic bronchitis, adult-onset asthma and acute exacerbations of asthma in adults, and small-scale studies have shown that prolonged antibiotic treatment was effective at greatly reducing the severity of the disease in some individuals (Ref 19-24).

In light of these results a protective vaccine against *C. pneumoniae* infection would be of considerable importance. There is not yet an effective vaccine for any human chlamydial infection. Nevertheless, studies with *C. trachomatis* and *C. psittaci* indicate that this is an attainable goal. For example, mice which have recovered from a lung infection with *C. trachomatis* are protected from infertility induced by a subsequent vaginal challenge (Ref 25). Similarly, sheep immunized with inactivated *C. psittaci* were protected from subsequent chlamydial-induced abortions and stillbirths (Ref 26). Protection from chlamydial infections has been associated with Th1 immune responses, particularly the induction of INF $\gamma$  - producing CD4+T-cells (Ref 27). The adoptive transfer of CD4+ cell lines or clones to nude or SCID mice conferred protection from challenge or cleared chronic disease (Ref 28,29), and in vivo depletion of CD4+ T cells exacerbated disease post-challenge (Ref 30,31). However, the presence of sufficiently high titres of neutralising antibody at mucosal surfaces can also exert a protective effect (Ref 32).

The extent of antigenic variation within the species *C. pneumoniae* is not well characterised. Serovars of *C. trachomatis* are defined on the basis of antigenic variation in MOMP, but published *C. pneumoniae* MOMP gene sequences show no variation between several diverse isolates of the organism (Ref 33-35). Regions of the protein known to be conserved in other chlamydial MOMPs are conserved in *C. pneumoniae* (Ref 33,34). One study has described a strain of *C. pneumoniae* with a MOMP of greater than usual molecular weight, but the gene for this has not been sequenced (Ref 1). Partial sequences of outer membrane protein 2 from nine diverse isolates were also found to be invariant (Ref 16). The genes for

HSP60 and HSP70 show little variation from other chlamydial species, as would be expected. The gene encoding a 76kDa antigen has been cloned from a single strain of *C. pneumoniae*. It has no significant similarity with other known chlamydial genes (Ref 4).

Many antigens recognised by immune sera to *C. pneumoniae* are conserved across all chlamydiae, but 98kDa, 76 kDa and 54 kDa proteins may be *C. pneumoniae*-specific (Ref 2, 4, 36). Immunoblotting of isolates with sera from patients does show variation of blotting patterns between isolates, indicating that serotypes *C. pneumoniae* may exist (Ref 1,16). However, the results are potentially confounded by the infection status of the patients, since immunoblot profiles of a patient's sera change with time post-infection. An assessment of the number and relative frequency of any serotypes, and the defining antigens, is not yet possible.

*C. pneumoniae* infection usually presents as an acute respiratory disease (i.e., cough, sore throat, hoarseness, and fever; abnormal chest sounds on auscultation). For most patients, the cough persists for 2 to 6 weeks, and recovery is slow. In approximately 10% of these cases, upper respiratory tract infection is followed by bronchitis or pneumonia. Furthermore, during a *C. pneumoniae* epidemic, subsequent co-infection with pneumococcus has been noted in about half of these pneumonia patients, particularly in the infirm and the elderly. As noted above, there is more and more evidence that *C. pneumoniae* infection is also linked to diseases other than respiratory infections.

The reservoir for the organism is presumably people. In contrast to *C. psittaci* infections, there is no known bird or animal reservoir. Transmission has not been clearly defined. It may result from direct contact with secretions, from fomites, or from airborne spread. There is a long incubation period, which may last for many months. Based on analysis of epidemics, *C. pneumoniae* appears to spread slowly through a population (case-to-case interval averaging 30 days) because infected persons are inefficient transmitters of the organism. Susceptibility to *C. pneumoniae* is universal. Reinfections occur during adulthood, following the primary infection as a child. *C. pneumoniae* appears to be an endemic disease throughout the world, noteworthy for superimposed intervals of increased incidence (epidemics) that persist for 2 to 3 years. *C. trachomatis* infection does not confer cross-immunity to *C. pneumoniae*. Infections are easily treated with oral antibiotics, tetracycline or erythromycin (2 g/d, for at least 10 to 14 d). A recently developed drug, azithromycin, is highly effective as a single-dose therapy against chlamydial infections.

In most instances, *C. pneumoniae* infection is often mild and without complications, and up to 90% of infections are subacute or unrecognized. Among children in industrialized countries, infections have been thought to be rare up to the age of 5 y, although a recent study (E Normann et al, *Chlamydia pneumoniae* in children with acute respiratory tract infections, *Acta Paediatrica*, 1998, Vol 87, Iss 1, pp 23-27) has reported that many children in this age group show PCR evidence of infection despite being seronegative, and estimates a prevalence of 17-19% in 2-4 y olds. In developing countries, the seroprevalence of *C. pneumoniae* antibodies among young children is elevated, and there are suspicions that *C. pneumoniae* may be an important cause of acute lower respiratory tract disease and mortality for infants and children in tropical regions of the world.

From seroprevalence studies and studies of local epidemics, the initial *C. pneumoniae* infection usually happens between the ages of 5 and 20 y. In the USA, for example, there are estimated to be 30,000 cases of childhood pneumonia each year caused by *C. pneumoniae*. Infections may cluster among groups of children or young adults (e.g., school pupils or military conscripts).

*C. pneumoniae* causes 10 to 25% of community-acquired lower respiratory tract infections (as reported from Sweden, Italy, Finland, and the USA). During an epidemic, *C. pneumoniae* infection may account for 50 to 60% of the cases of pneumonia. During these periods, also, more episodes of mixed infections with *S. pneumoniae* have been reported. Reinfection during adulthood is common; the clinical presentation tends to be milder. Based on population seroprevalence studies, there tends to be increased exposure with age, which is particularly evident among men. Some investigators have speculated that a persistent, asymptomatic *C. pneumoniae* infection state is common.

In adults of middle age or older, *C. pneumoniae* infection may progress to chronic bronchitis and sinusitis. A study in the USA revealed that the incidence of pneumonia caused by *C. pneumoniae* in persons younger than 60 years is 1 case per 1,000 persons per year; but in the elderly, the disease incidence rose three-fold. *C. pneumoniae* infection rarely leads to hospitalization, except in patients with an underlying illness.

#### SUMMARY OF THE INVENTION

The present invention provides purified and isolated DNA molecules that encode *Chlamydia* polypeptides designated CPN100634 (SEQ ID No: 1,2), which can be used in

methods to prevent, treat, and diagnose *Chlamydia* infection. The encoded polypeptides include polypeptides having the amino acid sequence shown in SEQ ID No:3. Those skilled in the art will appreciate that the invention also includes DNA molecules that encode mutants and derivatives of such polypeptides, which result from the addition, deletion, or substitution of non-essential amino acids as described herein. The invention also includes RNA molecules corresponding to the DNA molecules of the invention.

In addition to the DNA and RNA molecules, the invention includes the corresponding polypeptides and monospecific antibodies that specifically bind to such polypeptides.

The present invention has wide application and includes expression cassettes, vectors, and cells transformed or transfected with the polynucleotides of the invention. Accordingly, the present invention provides (i) a method for producing a polypeptide of the invention in a recombinant host system and related expression cassettes, vectors, and transformed or transfected cells; (ii) a live vaccine vector, such as a pox virus, *Salmonella typhimurium*, or *Vibrio cholerae* vector, containing a polynucleotide of the invention, such vaccine vectors being useful for, e.g., preventing and treating *Chlamydia* infection, in combination with a diluent or carrier, and related pharmaceutical compositions and associated therapeutic and/or prophylactic methods; (iii) a therapeutic and/or prophylactic method involving administration of an RNA or DNA molecule of the invention, either in a naked form or formulated with a delivery vehicle, a polypeptide or combination of polypeptides, or a monospecific antibody of the invention, and related pharmaceutical compositions; (iv) a method for diagnosing the presence of *Chlamydia* in a biological sample, which can involve the use of a DNA or RNA molecule, a monospecific antibody, or a polypeptide of the invention; and (v) a method for purifying a polypeptide of the invention by antibody-based affinity chromatography.

#### BRIEF DESCRIPTION OF THE DRAWINGS

The present invention will be further understood from the following description with reference to the drawings, in which:

Figure 1 shows the nucleotide sequence of the CPN100634 (SEQ ID No: 1 – entire sequence and SEQ ID No: 2 – coding sequence) and the deduced amino acid sequence of the CPN100634 protein from *Chlamydia pneumoniae* (SEQ ID No: 3).

Figure 2 shows the restriction enzyme analysis of the gene encoding the *C. pneumoniae* CPN100634 gene.

### DETAILED DESCRIPTION OF INVENTION

In the *C. pneumoniae* genome, open reading frames (ORFs) encoding chlamydial polypeptides have been identified. These polypeptides include polypeptides permanently found in the bacterial membrane structure, polypeptides that are present in the external vicinity of the bacterial membrane, include polypeptides permanently found in the inclusion membrane structure, polypeptides that are present in the external vicinity of the inclusion membrane, and polypeptides that are released into the cytoplasm of the infected cell. These polypeptides can be used in vaccination methods for preventing and treating *Chlamydia* infection.

According to a first aspect of the invention, there are provided isolated polynucleotides encoding the precursor and mature forms of *Chlamydia* polypeptides.

An isolated polynucleotide of the invention encodes (i) a polypeptide having an amino acid sequence that is homologous to a *Chlamydia* amino acid, the *Chlamydia* amino acid sequence being selected from the group consisting of:

(a) the amino acid sequences as shown: (SEQ ID No: 3)

The term "isolated polynucleotide" is defined as a polynucleotide removed from the environment in which it naturally occurs. For example, a naturally-occurring DNA molecule present in the genome of a living bacteria or as part of a gene bank is not isolated, but the same molecule separated from the remaining part of the bacterial genome, as a result of, *e.g.*, a cloning event (amplification), is isolated. Typically, an isolated DNA molecule is free from DNA regions (*e.g.*, coding regions) with which it is immediately contiguous at the 5' or 3' end, in the naturally occurring genome. Such isolated polynucleotides could be part of a vector or a composition and still be isolated in that such a vector or composition is not part of its natural environment.

A polynucleotide of the invention can be in the form of RNA or DNA (*e.g.*, cDNA, genomic DNA, or synthetic DNA), or modifications or combinations thereof. The DNA can be double-stranded or single-stranded, and, if single-stranded, can be the coding strand or the non-coding (anti-sense) strand. The sequence that encodes a polypeptide of the invention as shown in SEQ ID NOs: 1 and 2, can be (a) the coding sequence as shown in SEQ ID NOs:2 (b) a ribonucleotide sequence derived by transcription of (a); or (c) a different coding sequence; this latter, as a result of the redundancy or degeneracy of the genetic code, encodes the same



polypeptides as the DNA molecules of which the nucleotide sequences are illustrated in SEQ ID NOs:1 to 2.

By "polypeptide" or "protein" is meant any chain of amino acids, regardless of length or post-translational modification (*e.g.*, glycosylation or phosphorylation). Both terms are used interchangeably in the present application.

By "homologous amino acid sequence" is meant an amino acid sequence that differs from an amino acid sequence shown in SEQ ID No: 3, only by one or more conservative amino acid substitutions, or by one or more non-conservative amino acid substitutions, deletions, or additions located at positions at which they do not destroy the specific antigenicity of the polypeptide.

Preferably, such a sequence is at least 75%, more preferably 80%, and most preferably 90% identical to an amino acid sequence shown in SEQ ID No: 3.

Homologous amino acid sequences include sequences that are identical or substantially identical to an amino acid sequence as shown in SEQ ID No:3. By "amino acid sequence substantially identical" is meant a sequence that is at least 90%, preferably 95%, more preferably 97%, and most preferably 99% identical to an amino acid sequence of reference and that preferably differs from the sequence of reference, if at all, by a majority of conservative amino acid substitutions.

Conservative amino acid substitutions typically include substitutions among amino acids of the same class. These classes include, for example, amino acids having uncharged polar side chains, such as asparagine, glutamine, serine, threonine, and tyrosine; amino acids having basic side chains, such as lysine, arginine, and histidine; amino acids having acidic side chains, such as aspartic acid and glutamic acid; and amino acids having nonpolar side chains, such as glycine, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan, and cysteine.

Homology is typically measured using sequence analysis software (*e.g.*, Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, WI 53705). Similar amino acid sequences are aligned to obtain the maximum degree of homology (*i.e.*, identity). To this end, it may be necessary to artificially introduce gaps into the sequence. Once the optimal alignment has been set up, the degree of homology (*i.e.*, identity) is established by recording all of the

positions in which the amino acids of both sequences are identical, relative to the total number of positions.

Homologous polynucleotide sequences are defined in a similar way. Preferably, a homologous sequence is one that is at least 45%, more preferably 60%, and most preferably 85% identical to (i) a coding sequence of SEQ ID NOs:1 and 2.

Polypeptides having a sequence homologous to one of the sequences shown in SEQ ID NO: 3, include naturally-occurring allelic variants, as well as mutants or any other non-naturally occurring variants that are analogous in terms of antigenicity, to a polypeptide having a sequence as shown in SEQ ID NO: 3.

As is known in the art, an allelic variant is an alternate form of a polypeptide that is characterized as having a substitution, deletion, or addition of one or more amino acids that does not alter the biological function of the polypeptide. By "biological function" is meant the function of the polypeptide in the cells in which it naturally occurs, even if the function is not necessary for the growth or survival of the cells. For example, the biological function of a porin is to allow the entry into cells of compounds present in the extracellular medium. The biological function is distinct from the antigenic function. A polypeptide can have more than one biological function.

Allelic variants are very common in nature. For example, a bacterial species, *e.g.*, *C. pneumoniae*, is usually represented by a variety of strains that differ from each other by minor allelic variations. Indeed, a polypeptide that fulfills the same biological function in different strains can have an amino acid sequence that is not identical in each of the strains. Such an allelic variation may be equally reflected at the polynucleotide level.

Support for the use of allelic variants of polypeptide antigens comes from, *e.g.*, studies of the *Chlamydial* MOMP antigen. The amino acid sequence of the MOMP varies from strain to strain, yet cross-strain antibody binding plus neutralization of infectivity occurs, indicating that the MOMP, when used as an immunogen, is tolerant of amino acid variations.

Polynucleotides, *e.g.*, DNA molecules, encoding allelic variants can easily be retrieved by polymerase chain reaction (PCR) amplification of genomic bacterial DNA extracted by conventional methods. This involves the use of synthetic oligonucleotide primers matching upstream and downstream of the 5' and 3' ends of the encoding domain. Suitable primers can be designed according to the nucleotide sequence information provided in SEQ ID NOs:1 and 2.

Typically, a primer can consist of 10 to 40, preferably 15 to 25 nucleotides. It may be also advantageous to select primers containing C and G nucleotides in a proportion sufficient to ensure efficient hybridization; e.g., an amount of C and G nucleotides of at least 40%, preferably 50% of the total nucleotide amount.

5 Useful homologs that do not naturally occur can be designed using known methods for identifying regions of an antigen that are likely to be tolerant of amino acid sequence changes and/or deletions. For example, sequences of the antigen from different species can be compared to identify conserved sequences.

Polypeptide derivatives that are encoded by polynucleotides of the invention include,  
10 e.g., fragments, polypeptides having large internal deletions derived from full-length polypeptides, and fusion proteins.

Polypeptide fragments of the invention can be derived from a polypeptide having a sequence homologous to any of the sequences shown in SEQ ID NO: 3, to the extent that the fragments retain the substantial antigenicity of the parent polypeptide (specific antigenicity).  
15 Polypeptide derivatives can also be constructed by large internal deletions that remove a substantial part of the parent polypeptide, while retaining specific antigenicity. Generally, polypeptide derivatives should be about at least 12 amino acids in length to maintain antigenicity. Advantageously, they can be at least 20 amino acids, preferably at least 50 amino acids, more preferably at least 75 amino acids, and most preferably at least 100 amino acids in  
20 length.

Useful polypeptide derivatives, e.g., polypeptide fragments, can be designed using computer-assisted analysis of amino acid sequences in order to identify sites in protein antigens having potential as surface-exposed, antigenic regions (Ref 37).

Polypeptide fragments and polypeptides having large internal deletions can be used  
25 for revealing epitopes that are otherwise masked in the parent polypeptide and that may be of importance for inducing a protective T cell-dependent immune response. Deletions can also remove immunodominant regions of high variability among strains.

It is an accepted practice in the field of immunology to use fragments and variants of protein immunogens as vaccines, as all that is required to induce an immune response to a  
30 protein is a small (e.g., 8 to 10 amino acid) immunogenic region of the protein. This has been done for a number of vaccines against pathogens other than *Chlamydia*. For example, short

synthetic peptides corresponding to surface-exposed antigens of pathogens such as murine mammary tumor virus, peptide containing 11 amino acids; (Ref 38), Semliki Forest virus, peptide containing 16 amino acids (Ref 39), and canine parvovirus, 2 overlapping peptides, each containing 15 amino acids (Ref 40), have been shown to be effective vaccine antigens against  
5 their respective pathogens.

Polynucleotides encoding polypeptide fragments and polypeptides having large internal deletions can be constructed using standard methods (Ref 41), for example, by PCR, including inverse PCR, by restriction enzyme treatment of the cloned DNA molecules, or by the method of Kunkel *et al.* (Ref 42) biological material available at Stratagene.

10 A polypeptide derivative can also be produced as a fusion polypeptide that contains a polypeptide or a polypeptide derivative of the invention fused, *e.g.*, at the N- or C-terminal end, to any other polypeptide (hereinafter referred to as a peptide tail). Such a product can be easily obtained by translation of a genetic fusion, *i.e.*, a hybrid gene. Vectors for expressing fusion polypeptides are commercially available, such as the pMal-c2 or pMal-p2 systems of New  
15 England Biolabs, in which the peptide tail is a maltose binding protein, the glutathione-S-transferase system of Pharmacia, or the His-Tag system available from Novagen. These and other expression systems provide convenient means for further purification of polypeptides and derivatives of the invention.

Another particular example of fusion polypeptides included in invention includes a  
20 polypeptide or polypeptide derivative of the invention fused to a polypeptide having adjuvant activity, such as, *e.g.*, subunit B of either cholera toxin or *E. coli* heat-labile toxin. Several possibilities are can be used for achieving fusion. First, the polypeptide of the invention can be fused to the N-, or preferably, to the C-terminal end of the polypeptide having adjuvant activity. Second, a polypeptide fragment of the invention can be fused within the amino acid sequence of  
25 the polypeptide having adjuvant activity.

As stated above, the polynucleotides of the invention encode *Chlamydia* polypeptides in precursor or mature form. They can also encode hybrid precursors containing heterologous signal peptides, which can mature into polypeptides of the invention. By "heterologous signal peptide" is meant a signal peptide that is not found in the naturally-occurring precursor of a  
30 polypeptide of the invention.

A polynucleotide of the invention, having a homologous coding sequence, hybridizes, preferably under stringent conditions, to a polynucleotide having a sequence as shown in SEQ ID NOs:1 to 2. Hybridization procedures are, *e.g.*, described in Ausubel *et al.*, (Ref 41), Silhavy *et al.* (Ref 43); Davis *et al.* (ref 44). Important parameters that can be considered for optimizing hybridization conditions are reflected in a formula that allows calculation of a critical value, the melting temperature above which two complementary DNA strands separate from each other Ref 45). This formula is as follows:  $T_m = 81.5 + 0.5 \times (\% \text{ G+C}) + 1.6 \log (\text{positive ion concentration}) - 0.6 \times (\% \text{ formamide})$ . Under appropriate stringency conditions, hybridization temperature ( $T_h$ ) is approximately 20 to 40°C, 20 to 25°C, or, preferably 30 to 40°C below the calculated  $T_m$ . Those skilled in the art will understand that optimal temperature and salt conditions can be readily determined empirically in preliminary experiments using conventional procedures.

For example, stringent conditions can be achieved, both for pre-hybridizing and hybridizing incubations, (i) within 4-16 hours at 42°C, in 6 x SSC containing 50% formamide or (ii) within 4-16 hours at 65°C in an aqueous 6 x SSC solution (1 M NaCl, 0.1 M sodium citrate (pH 7.0)).

For polynucleotides containing 30 to 600 nucleotides, the above formula is used and then is corrected by subtracting (600/polynucleotide size in base pairs). Stringency conditions are defined by a  $T_h$  that is 5 to 10°C below  $T_m$ .

Hybridization conditions with oligonucleotides shorter than 20-30 bases do not exactly follow the rules set forth above. In such cases, the formula for calculating the  $T_m$  is as follows:  $T_m = 4 \times (\text{G+C}) + 2 \times (\text{A+T})$ . For example, an 18 nucleotide fragment of 50% G+C would have an approximate  $T_m$  of 54°C.

A polynucleotide molecule of the invention, containing RNA, DNA, or modifications or combinations thereof, can have various applications. For example, a DNA molecule can be used (i) in a process for producing the encoded polypeptide in a recombinant host system, (ii) in the construction of vaccine vectors such as poxviruses, which are further used in methods and compositions for preventing and/or treating *Chlamydia* infection, (iii) as a vaccine agent (as well as an RNA molecule), in a naked form or formulated with a delivery vehicle and, (iv) in the construction of attenuated *Chlamydia* strains that can over-express a polynucleotide of the invention or express it in a non-toxic, mutated form.

According to a second aspect of the invention, there is therefore provided (i) an expression cassette containing a DNA molecule of the invention placed under the control of the elements required for expression, in particular under the control of an appropriate promoter; (ii) an expression vector containing an expression cassette of the invention; (iii) a procaryotic or eucaryotic cell transformed or transfected with an expression cassette and/or vector of the invention, as well as (iv) a process for producing a polypeptide or polypeptide derivative encoded by a polynucleotide of the invention, which involves culturing a procaryotic or eucaryotic cell transformed or transfected with an expression cassette and/or vector of the invention, under conditions that allow expression of the DNA molecule of the invention and, recovering the encoded polypeptide or polypeptide derivative from the cell culture.

A recombinant expression system can be selected from procaryotic and eucaryotic hosts. Eucaryotic hosts include yeast cells (*e.g.*, *Saccharomyces cerevisiae* or *Pichia pastoris*), mammalian cells (*e.g.*, COS1, NIH3T3, or JEG3 cells), arthropods cells (*e.g.*, *Spodoptera frugiperda* (SF9) cells), and plant cells. Preferably, a procaryotic host such as *E. coli* is used. Bacterial and eucaryotic cells are available from a number of different sources to those skilled in the art, *e.g.*, the American Type Culture Collection (ATCC; Rockville, Maryland).

The choice of the expression system depends on the features desired for the expressed polypeptide. For example, it may be useful to produce a polypeptide of the invention in a particular lipidated form or any other form.

The choice of the expression cassette will depend on the host system selected as well as the features desired for the expressed polypeptide. Typically, an expression cassette includes a promoter that is functional in the selected host system and can be constitutive or inducible; a ribosome binding site; a start codon (ATG) if necessary, a region encoding a signal peptide, *e.g.*, a lipidation signal peptide; a DNA molecule of the invention; a stop codon; and optionally a 3' terminal region (translation and/or transcription terminator). The signal peptide encoding region is adjacent to the polynucleotide of the invention and placed in proper reading frame. The signal peptide-encoding region can be homologous or heterologous to the DNA molecule encoding the mature polypeptide and can be specific to the secretion apparatus of the host used for expression. The open reading frame constituted by the DNA molecule of the invention, solely or together with the signal peptide, is placed under the control of the promoter so that transcription and translation occur in the host system. Promoters, signal peptide encoding regions are widely

known and available to those skilled in the art and includes, for example, the promoter of *Salmonella typhimurium* (and derivatives) that is inducible by arabinose (promoter araB) and is functional in Gram-negative bacteria such as *E. coli* (as described in U.S. Patent No. 5,028,530 and in Cagnon *et al.*, (Ref 46); the promoter of the gene of bacteriophage T7 encoding RNA polymerase, that is functional in a number of *E. coli* strains expressing T7 polymerase (described in U.S. Patent No. 4,952,496); OspA lipidation signal peptide ; and RlpB lipidation signal peptide (Ref 47).

The expression cassette is typically part of an expression vector, which is selected for its ability to replicate in the chosen expression system. Expression vectors (*e.g.*, plasmids or viral vectors) can be chosen from those described in Pouwels *et al.* (Cloning Vectors: A Laboratory Manual 1985, Supp. 1987). They can be purchased from various commercial sources.

Methods for transforming/transfecting host cells with expression vectors will depend on the host system selected as described in Ausubel *et al.*, (Ref 41).

Upon expression, a recombinant polypeptide of the invention (or a polypeptide derivative) is produced and remains in the intracellular compartment, is secreted/excreted in the extracellular medium or in the periplasmic space, or is embedded in the cellular membrane. The polypeptide can then be recovered in a substantially purified form from the cell extract or from the supernatant after centrifugation of the recombinant cell culture. Typically, the recombinant polypeptide can be purified by antibody-based affinity purification or by any other method that can be readily adapted by a person skilled in the art, such as by genetic fusion to a small affinity binding domain. Antibody-based affinity purification methods are also available for purifying a polypeptide of the invention extracted from a *Chlamydia* strain. Antibodies useful for purifying by immunoaffinity the polypeptides of the invention can be obtained as described below.

A polynucleotide of the invention can also be useful in the vaccine field, *e.g.*, for achieving DNA vaccination. There are two major possibilities, either using a viral or bacterial host as gene delivery vehicle (live vaccine vector) or administering the gene in a free form, *e.g.*, inserted into a plasmid. Therapeutic or prophylactic efficacy of a polynucleotide of the invention can be evaluated as described below.

Accordingly, in a third aspect of the invention, there is provided (i) a vaccine vector such as a poxvirus, containing a DNA molecule of the invention, placed under the control of

elements required for expression; (ii) a composition of matter containing a vaccine vector of the invention, together with a diluent or carrier; particularly, (iii) a pharmaceutical composition containing a therapeutically or prophylactically effective amount of a vaccine vector of the invention; (iv) a method for inducing an immune response against *Chlamydia* in a mammal (e.g., a human; alternatively, the method can be used in veterinary applications for treating or preventing *Chlamydia* infection of animals, e.g., cats or birds), which involves administering to the mammal an immunogenically effective amount of a vaccine vector of the invention to elicit an immune response, e.g., a protective or therapeutic immune response to *Chlamydia*; and particularly, (v) a method for preventing and/or treating a *Chlamydia* (e.g., *C. trachomatis*, *C. psittaci*, *C. pneumonia*, *C. pecorum*) infection, which involves administering a prophylactic or therapeutic amount of a vaccine vector of the invention to an individual in need. Additionally, the third aspect of the invention encompasses the use of a vaccine vector of the invention in the preparation of a medicament for preventing and/or treating *Chlamydia* infection.

A vaccine vector of the invention can express one or several polypeptides or derivatives of the invention, as well as at least one additional *Chlamydia* antigen, fragment, homolog, mutant, or derivative thereof. In addition, it can express a cytokine, such as interleukin-2 (IL-2) or interleukin-12 (IL-12), that enhances the immune response (adjuvant effect). Thus, a vaccine vector can include an additional DNA sequence encoding, e.g., a chlamydial antigen, or a cytokine, placed under the control of elements required for expression in a mammalian cell.

Alternatively, a composition of the invention can include several vaccine vectors, each of them being capable of expressing a polypeptide or derivative of the invention. A composition can also contain a vaccine vector capable of expressing an additional *Chlamydia* antigen, or a subunit, fragment, homolog, mutant, or derivative thereof; or a cytokine such as IL-2 or IL-12.

In vaccination methods for treating or preventing infection in a mammal, a vaccine vector of the invention can be administered by any conventional route in use in the vaccine field, particularly, to a mucosal (e.g., ocular, intranasal, oral, gastric, pulmonary, intestinal, rectal, vaginal, or urinary tract) surface or via the parenteral (e.g., subcutaneous, intradermal, intramuscular, intravenous, or intraperitoneal) route. Preferred routes depend upon the choice of the vaccine vector. The administration can be achieved in a single dose or repeated at intervals.



The appropriate dosage depends on various parameters understood by skilled artisans such as the vaccine vector itself, the route of administration or the condition of the mammal to be vaccinated (weight, age and the like).

Live vaccine vectors available in the art include viral vectors such as adenoviruses  
5 and poxviruses as well as bacterial vectors, *e.g.*, *Shigella*, *Salmonella*, *Vibrio cholerae*,  
*Lactobacillus*, Bacille bilié de Calmette-Guérin (BCG), and *Streptococcus*.

An example of an adenovirus vector, as well as a method for constructing an  
adenovirus vector capable of expressing a DNA molecule of the invention, are described in  
U.S. Patent No. 4,920,209. Poxvirus vectors that can be used include, *e.g.*, vaccinia and canary  
10 pox virus, described in U.S. Patent No. 4,722,848 and U.S. Patent No. 5,364,773, respectively  
(also see, *e.g.*, Tartaglia *et al.*, Virology (1992) 188:217) for a description of a vaccinia virus  
vector; and Taylor *et al.*, Vaccine (1995) 13:539 for a reference of a canary pox). Poxvirus  
vectors capable of expressing a polynucleotide of the invention can be obtained by homologous  
recombination as described in Kieny *et al.*, Nature (1984) 312:163 so that the polynucleotide of  
15 the invention is inserted in the viral genome under appropriate conditions for expression in  
mammalian cells. Generally, the dose of vaccine viral vector, for therapeutic or prophylactic use,  
can be of from about  $1 \times 10^4$  to about  $1 \times 10^{11}$ , advantageously from about  $1 \times 10^7$  to about  $1 \times 10^{10}$ ,  
preferably of from about  $1 \times 10^7$  to about  $1 \times 10^9$  plaque-forming units per kilogram. Preferably,  
viral vectors are administered parenterally; for example, in 3 doses, 4 weeks apart. Those skilled  
20 in the art recognize that it is preferable to avoid adding a chemical adjuvant to a composition  
containing a viral vector of the invention and thereby minimizing the immune response to the  
viral vector itself.

Non-toxicogenic *Vibrio cholerae* mutant strains that are useful as a live oral vaccine  
are described in Mekalanos *et al.*, Nature (1983) 306:551 and U.S. Patent No. 4,882,278 (strain  
25 in which a substantial amount of the coding sequence of each of the two *ctxA* alleles has been  
deleted so that no functional *cholerae* toxin is produced); WO 92/11354 (strain in which the *irgA*  
locus is inactivated by mutation; this mutation can be combined in a single strain with *ctxA*  
mutations); and WO 94/1533 (deletion mutant lacking functional *ctxA* and *attRS1* DNA  
sequences). These strains can be genetically engineered to express heterologous antigens, as  
30 described in WO 94/19482. An effective vaccine dose of a *Vibrio cholerae* strain capable of  
expressing a polypeptide or polypeptide derivative encoded by a DNA molecule of the invention

can contain, e.g., about  $1 \times 10^5$  to about  $1 \times 10^9$ , preferably about  $1 \times 10^6$  to about  $1 \times 10^8$  viable bacteria in an appropriate volume for the selected route of administration. Preferred routes of administration include all mucosal routes; most preferably, these vectors are administered intranasally or orally.

5           Attenuated *Salmonella typhimurium* strains, genetically engineered for recombinant expression of heterologous antigens or not, and their use as oral vaccines are described in Nakayama *et al.* (Bio/Technology (1988) 6:693) and WO 92/11361. Preferred routes of administration include all mucosal routes; most preferably, these vectors are administered intranasally or orally.

10           Others bacterial strains useful as vaccine vectors are described in High *et al.*, EMBO (1992) 11:1991 and Sizemore *et al.*, Science (1995) 270:299 (*Shigella flexneri*); Medaglini *et al.*, Proc. Natl. Acad. Sci. USA (1995) 92:6868 (*Streptococcus gordonii*); and Flynn J.L., Cell. Mol. Biol. (1994) 40 (suppl. I):31, WO 88/6626, WO 90/0594, WO 91/13157, WO 92/1796, and WO 92/21376 (Bacille Calmette Guerin).

15           In bacterial vectors, polynucleotide of the invention can be inserted into the bacterial genome or can remain in a free state, carried on a plasmid.

          An adjuvant can also be added to a composition containing a vaccine bacterial vector. A number of adjuvants are known to those skilled in the art. Preferred adjuvants can be selected from the list provided below.

20           According to a fourth aspect of the invention, there is also provided (i) a composition of matter containing a polynucleotide of the invention, together with a diluent or carrier; (ii) a pharmaceutical composition containing a therapeutically or prophylactically effective amount of a polynucleotide of the invention; (iii) a method for inducing an immune response against *Chlamydia*, in a mammal, by administering to the mammal, an immunogenically effective  
25           amount of a polynucleotide of the invention to elicit an immune response, e.g., a protective immune response to *Chlamydia*; and particularly, (iv) a method for preventing and/or treating a *Chlamydia* (e.g., *C. trachomatis*, *C. psittaci*, *C. pneumoniae*, or *C. pecorum*) infection, by administering a prophylactic or therapeutic amount of a polynucleotide of the invention to an individual in need. Additionally, the fourth aspect of the invention encompasses the use of a  
30           polynucleotide of the invention in the preparation of a medicament for preventing and/or treating *Chlamydia* infection. The fourth aspect of the invention preferably includes the use of a DNA

molecule placed under conditions for expression in a mammalian cell, *e.g.*, in a plasmid that is unable to replicate in mammalian cells and to substantially integrate in a mammalian genome.

Polynucleotides (DNA or RNA) of the invention can also be administered as such to a mammal for vaccine, *e.g.*, therapeutic or prophylactic, purpose. When a DNA molecule of the invention is used, it can be in the form of a plasmid that is unable to replicate in a mammalian cell and unable to integrate in the mammalian genome. Typically, a DNA molecule is placed under the control of a promoter suitable for expression in a mammalian cell. The promoter can function ubiquitously or tissue-specifically. Examples of non-tissue specific promoters include the early Cytomegalovirus (CMV) promoter (described in U.S. Patent No. 4,168,062) and the Rous Sarcoma Virus promoter (described in Norton & Coffin, *Molec. Cell Biol.* (1985) 5:281). The desmin promoter (Li *et al.*, *Gene* (1989) 78:243, Li & Paulin, *J. Biol. Chem.* (1991) 266:6562 and Li & Paulin, *J. Biol. Chem.* (1993) 268:10403) is tissue-specific and drives expression in muscle cells. More generally, useful vectors are described, *i.a.*, WO 94/21797 and Hartikka *et al.*, *Human Gene Therapy* (1996) 7:1205.

For DNA/RNA vaccination, the polynucleotide of the invention can encode a precursor or a mature form. When it encodes a precursor form, the precursor form can be homologous or heterologous. In the latter case, a eucaryotic leader sequence can be used, such as the leader sequence of the tissue-type plasminogen factor (tPA).

A composition of the invention can contain one or several polynucleotides of the invention. It can also contain at least one additional polynucleotide encoding another *Chlamydia* antigen such as urease subunit A, B, or both; or a fragment, derivative, mutant, or analog thereof. A polynucleotide encoding a cytokine, such as interleukin-2 (IL-2) or interleukin-12 (IL-12), can also be added to the composition so that the immune response is enhanced. These additional polynucleotides are placed under appropriate control for expression. Advantageously, DNA molecules of the invention and/or additional DNA molecules to be included in the same composition, can be carried in the same plasmid.

Standard techniques of molecular biology for preparing and purifying polynucleotides can be used in the preparation of polynucleotide therapeutics of the invention. For use as a vaccine, a polynucleotide of the invention can be formulated according to various methods.

First, a polynucleotide can be used in a naked form, free of any delivery vehicles, such as anionic liposomes, cationic lipids, microparticles, *e.g.*, gold microparticles, precipitating

agents, *e.g.*, calcium phosphate, or any other transfection-facilitating agent. In this case, the polynucleotide can be simply diluted in a physiologically acceptable solution, such as sterile saline or sterile buffered saline, with or without a carrier. When present, the carrier preferably is isotonic, hypotonic, or weakly hypertonic, and has a relatively low ionic strength, such as provided by a sucrose solution, *e.g.*, a solution containing 20% sucrose.

Alternatively, a polynucleotide can be associated with agents that assist in cellular uptake. It can be, *i.a.*, (i) complemented with a chemical agent that modifies the cellular permeability, such as bupivacaine (see, *e.g.*, WO 94/16737), (ii) encapsulated into liposomes, or (iii) associated with cationic lipids or silica, gold, or tungsten microparticles.

Anionic and neutral liposomes are well-known in the art (see, *e.g.*, Liposomes: A Practical Approach, RPC New Ed, IRL press (1990), for a detailed description of methods for making liposomes) and are useful for delivering a large range of products, including polynucleotides.

Cationic lipids are also known in the art and are commonly used for gene delivery. Such lipids include Lipofectin™ also known as DOTMA (N-[1-(2,3-dioleoyloxy)propyl]-N,N,N-trimethylammonium chloride), DOTAP (1,2-bis(oleoyloxy)-3-(trimethylammonio)propane), DDAB (dimethyldioctadecylammonium bromide), DOGS (dioctadecylamidologlycyl spermine) and cholesterol derivatives such as DC-Chol (3 beta-(N-(N',N'-dimethyl aminomethane)-carbamoyl) cholesterol). A description of these cationic lipids can be found in EP 187,702, WO 90/11092, U.S. Patent No. 5,283,185, WO 91/15501, WO 95/26356, and U.S. Patent No. 5,527,928. Cationic lipids for gene delivery are preferably used in association with a neutral lipid such as DOPE (dioleoyl phosphatidylethanolamine), as, for example, described in WO 90/11092.

Other transfection-facilitating compounds can be added to a formulation containing cationic liposomes. A number of them are described in, *e.g.*, WO 93/18759, WO 93/19768, WO 94/25608, and WO 95/2397. They include, *i.a.*, spermine derivatives useful for facilitating the transport of DNA through the nuclear membrane (see, for example, WO 93/18759) and membrane-permeabilizing compounds such as GALA, Gramicidine S, and cationic bile salts (see, for example, WO 93/19768).

Gold or tungsten microparticles can also be used for gene delivery, as described in WO 91/359, WO 93/17706, and Tang *et al.* (Nature (1992) 356:152). In this case, the

microparticle-coated polynucleotides can be injected *via* intradermal or intraepidermal routes using a needleless injection device ("gene gun"), such as those described in U.S. Patent No. 4,945,050, U.S. Patent No. 5,015,580, and WO 94/24263.

5 The amount of DNA to be used in a vaccine recipient depends, *e.g.*, on the strength of the promoter used in the DNA construct, the immunogenicity of the expressed gene product, the condition of the mammal intended for administration (*e.g.*, the weight, age, and general health of the mammal), the mode of administration, and the type of formulation. In general, a therapeutically or prophylactically effective dose from about 1  $\mu$ g to about 1 mg, preferably, from about 10  $\mu$ g to about 800  $\mu$ g and, more preferably, from about 25  $\mu$ g to about 250  $\mu$ g, can be administered to human adults. The administration can be achieved in a single dose or repeated at intervals.

15 The route of administration can be any conventional route used in the vaccine field. As general guidance, a polynucleotide of the invention can be administered *via* a mucosal surface, *e.g.*, an ocular, intranasal, pulmonary, oral, intestinal, rectal, vaginal, and urinary tract surface; or *via* a parenteral route, *e.g.*, by an intravenous, subcutaneous, intraperitoneal, intradermal, intraepidermal, or intramuscular route. The choice of the administration route will depend on, *e.g.*, the formulation that is selected. A polynucleotide formulated in association with bupivacaine is advantageously administered into muscles. When a neutral or anionic liposome or a cationic lipid, such as DOTMA or DC-Chol, is used, the formulation can be advantageously injected *via* intravenous, intranasal (aerosolization), intramuscular, intradermal, and subcutaneous routes. A polynucleotide in a naked form can advantageously be administered *via* the intramuscular, intradermal, or sub-cutaneous routes.

25 Although not absolutely required, such a composition can also contain an adjuvant. If so, a systemic adjuvant that does not require concomitant administration in order to exhibit an adjuvant effect is preferable such as, *e.g.*, QS21, which is described in U.S. Patent No. 5,057,546.

The sequence information provided in the present application enables the design of specific nucleotide probes and primers that can be useful in diagnosis. Accordingly, in a fifth aspect of the invention, there is provided a nucleotide probe or primer having a sequence found in or derived by degeneracy of the genetic code from a sequence shown in SEQ ID NO:1 to 2.

30 The term "probe" as used in the present application refers to DNA (preferably single stranded) or RNA molecules (or modifications or combinations thereof) that hybridize under the

stringent conditions, as defined above, to nucleic acid molecules having sequences homologous to those shown in SEQ ID NOs:1 and 2, or to a complementary or anti-sense sequence. Generally, probes are significantly shorter than full-length sequences shown in SEQ ID NOs:1 and 2; for example, they can contain from about 5 to about 100, preferably from about 10 to about 80 nucleotides. In particular, probes have sequences that are at least 75%, preferably at least 85%, more preferably 95% homologous to a portion of a sequence as shown in SEQ ID NOs:1 and 2 or that are complementary to such sequences. Probes can contain modified bases such as inosine, methyl-5-deoxycytidine, deoxyuridine, dimethylamino-5-deoxyuridine, or diamino-2, 6-purine. Sugar or phosphate residues can also be modified or substituted. For example, a deoxyribose residue can be replaced by a polyamide (Nielsen *et al.*, Science (1991) 254:1497) and phosphate residues can be replaced by ester groups such as diphosphate, alkyl, arylphosphonate and phosphorothioate esters. In addition, the 2'-hydroxyl group on ribonucleotides can be modified by including, *e.g.*, alkyl groups.

Probes of the invention can be used in diagnostic tests, as capture or detection probes. Such capture probes can be conventionally immobilized on a solid support, directly or indirectly, by covalent means or by passive adsorption. A detection probe can be labelled by a detection marker selected from radioactive isotopes; enzymes such as peroxidase, alkaline phosphatase, and enzymes able to hydrolyze a chromogenic, fluorogenic, or luminescent substrate; compounds that are chromogenic, fluorogenic, or luminescent; nucleotide base analogs; and biotin.

Probes of the invention can be used in any conventional hybridization technique, such as dot blot (Maniatis *et al.*, Molecular Cloning: A Laboratory Manual (1982) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York), Southern blot (Southern, J. Mol. Biol. (1975) 98:503), northern blot (identical to Southern blot to the exception that RNA is used as a target), or the sandwich technique (Dunn *et al.*, Cell (1977) 12:23). The latter technique involves the use of a specific capture probe and/or a specific detection probe with nucleotide sequences that at least partially differ from each other.

A primer is usually a probe of about 10 to about 40 nucleotides that is used to initiate enzymatic polymerization of DNA in an amplification process (*e.g.*, PCR), in an elongation process, or in a reverse transcription method. In a diagnostic method involving PCR, primers can be labelled.

Thus, the invention also encompasses (i) a reagent containing a probe of the invention for detecting and/or identifying the presence of *Chlamydia* in a biological material; (ii) a method for detecting and/or identifying the presence of *Chlamydia* in a biological material, in which (a) a sample is recovered or derived from the biological material, (b) DNA or RNA is extracted from the material and denatured, and (c) exposed to a probe of the invention, for example, a capture, detection probe or both, under stringent hybridization conditions, such that hybridization is detected; and (iii) a method for detecting and/or identifying the presence of *Chlamydia* in a biological material, in which (a) a sample is recovered or derived from the biological material, (b) DNA is extracted therefrom, (c) the extracted DNA is primed with at least one, and preferably two, primers of the invention and amplified by polymerase chain reaction, and (d) the amplified DNA fragment is produced.

As previously mentioned, polypeptides that can be produced upon expression of the newly identified open reading frames are useful vaccine agents.

Therefore, a sixth aspect of the invention features a substantially purified polypeptide or polypeptide derivative having an amino acid sequence encoded by a polynucleotide of the invention.

A "substantially purified polypeptide" is defined as a polypeptide that is separated from the environment in which it naturally occurs and/or that is free of the majority of the polypeptides that are present in the environment in which it was synthesized. For example, a substantially purified polypeptide is free from cytoplasmic polypeptides. Those skilled in the art will understand that the polypeptides of the invention can be purified from a natural source, *i.e.*, a *Chlamydia* strain, or can be produced by recombinant means.

Homologous polypeptides or polypeptide derivatives encoded by polynucleotides of the invention can be screened for specific antigenicity by testing cross-reactivity with an antiserum raised against the polypeptide of reference having an amino acid sequence as shown in SEQ ID NOs:3. Briefly, a monospecific hyperimmune antiserum can be raised against a purified reference polypeptide as such or as a fusion polypeptide, for example, an expression product of MBP, GST, or His-tag systems or a synthetic peptide predicted to be antigenic. The homologous polypeptide or derivative screened for specific antigenicity can be produced as such or as a fusion polypeptide. In this latter case and if the antiserum is also raised against a fusion polypeptide, two different fusion systems are employed. Specific antigenicity can be determined

according to a number of methods, including Western blot (Towbin *et al.*, Proc. Natl. Acad. Sci. USA (1979) 76:4350), dot blot, and ELISA, as described below.

In a Western blot assay, the product to be screened, either as a purified preparation or a total *E. coli* extract, is submitted to SDS-Page electrophoresis as described by Laemmli (Nature (1970) 227:680). After transfer to a nitrocellulose membrane, the material is further incubated with the monospecific hyperimmune antiserum diluted in the range of dilutions from about 1:5 to about 1:5000, preferably from about 1:100 to about 1:500. Specific antigenicity is shown once a band corresponding to the product exhibits reactivity at any of the dilutions in the above range.

In an ELISA assay, the product to be screened is preferably used as the coating antigen. A purified preparation is preferred, although a whole cell extract can also be used. Briefly, about 100  $\mu$ l of a preparation at about 10  $\mu$ g protein/ml are distributed into wells of a 96-well polycarbonate ELISA plate. The plate is incubated for 2 hours at 37°C then overnight at 4°C. The plate is washed with phosphate buffer saline (PBS) containing 0.05% Tween 20 (PBS/Tween buffer). The wells are saturated with 250  $\mu$ l PBS containing 1% bovine serum albumin (BSA) to prevent non-specific antibody binding. After 1 hour incubation at 37°C, the plate is washed with PBS/Tween buffer. The antiserum is serially diluted in PBS/Tween buffer containing 0.5% BSA. 100  $\mu$ l of dilutions are added per well. The plate is incubated for 90 minutes at 37°C, washed and evaluated according to standard procedures. For example, a goat anti-rabbit peroxidase conjugate is added to the wells when specific antibodies were raised in rabbits. Incubation is carried out for 90 minutes at 37°C and the plate is washed. The reaction is developed with the appropriate substrate and the reaction is measured by colorimetry (absorbance measured spectrophotometrically). Under the above experimental conditions, a positive reaction is shown by O.D. values greater than a non immune control serum.

In a dot blot assay, a purified product is preferred, although a whole cell extract can also be used. Briefly, a solution of the product at about 100  $\mu$ g/ml is serially two-fold diluted in 50 mM Tris-HCl (pH 7.5). 100  $\mu$ l of each dilution are applied to a nitrocellulose membrane 0.45  $\mu$ m set in a 96-well dot blot apparatus (Biorad). The buffer is removed by applying vacuum to the system. Wells are washed by addition of 50 mM Tris-HCl (pH 7.5) and the membrane is air-dried. The membrane is saturated in blocking buffer (50 mM Tris-HCl (pH 7.5) 0.15 M NaCl, 10 g/L skim milk) and incubated with an antiserum dilution from about 1:50 to about 1:5000, preferably about 1:500. The reaction is revealed according to standard procedures. For example,



a goat anti-rabbit peroxidase conjugate is added to the wells when rabbit antibodies are used. Incubation is carried out 90 minutes at 37°C and the blot is washed. The reaction is developed with the appropriate substrate and stopped. The reaction is measured visually by the appearance of a colored spot, *e.g.*, by colorimetry. Under the above experimental conditions, a positive  
 5 reaction is shown once a colored spot is associated with a dilution of at least about 1:5, preferably of at least about 1:500.

Therapeutic or prophylactic efficacy of a polypeptide or derivative of the invention can be evaluated as described below.

According to a seventh aspect of the invention, there is provided (i) a composition of  
 10 matter containing a polypeptide of the invention together with a diluent or carrier; in particular, (ii) a pharmaceutical composition containing a therapeutically or prophylactically effective amount of a polypeptide of the invention; (iii) a method for inducing an immune response against  
 15 *Chlamydia* in a mammal, by administering to the mammal an immunogenically effective amount of a polypeptide of the invention to elicit an immune response, *e.g.*, a protective immune response to *Chlamydia*; and particularly, (iv) a method for preventing and/or treating a  
 20 *Chlamydia* (*e.g.*, *C. trachomatis*, *C. psittaci*, *C. pneumoniae*, or *C. pecorum*) infection, by administering a prophylactic or therapeutic amount of a polypeptide of the invention to an individual in need. Additionally, the seventh aspect of the invention encompasses the use of a polypeptide of the invention in the preparation of a medicament for preventing and/or treating  
 25 *Chlamydia* infection.

The immunogenic compositions of the invention can be administered by any conventional route in use in the vaccine field, in particular to a mucosal (*e.g.*, ocular, intranasal, pulmonary, oral, gastric, intestinal, rectal, vaginal, or urinary tract) surface or *via* the parenteral (*e.g.*, subcutaneous, intradermal, intramuscular, intravenous, or intraperitoneal) route. The  
 25 choice of the administration route depends upon a number of parameters, such as the adjuvant associated with the polypeptide. For example, if a mucosal adjuvant is used, the intranasal or oral route will be preferred and if a lipid formulation or an aluminum compound is used, the parenteral route will be preferred. In the latter case, the sub-cutaneous or intramuscular route is most preferred. The choice can also depend upon the nature of the vaccine agent. For example,  
 30 a polypeptide of the invention fused to CTB or LTB will be best administered to a mucosal surface.

A composition of the invention can contain one or several polypeptides or derivatives of the invention. It can also contain at least one additional *Chlamydia* antigen, or a subunit, fragment, homolog, mutant, or derivative thereof.

For use in a composition of the invention, a polypeptide or derivative thereof can be formulated into or with liposomes, preferably neutral or anionic liposomes, microspheres, ISCOMS, or virus-like-particles (VLPs) to facilitate delivery and/or enhance the immune response. These compounds are readily available to one skilled in the art; for example, see Liposomes: A Practical Approach (*supra*).

Adjuvants other than liposomes and the like can also be used and are known in the art. A appropriate selection can conventionally be made by those skilled in the art, for example, from the list provided below.

Administration can be achieved in a single dose or repeated as necessary at intervals as can be determined by one skilled in the art. For example, a priming dose can be followed by three booster doses at weekly or monthly intervals. An appropriate dose depends on various parameters including the recipient (*e.g.*, adult or infant), the particular vaccine antigen, the route and frequency of administration, the presence/absence or type of adjuvant, and the desired effect (*e.g.*, protection and/or treatment), as can be determined by one skilled in the art. In general, a vaccine antigen of the invention can be administered by a mucosal route in an amount from about 10  $\mu$ g to about 500 mg, preferably from about 1 mg to about 200 mg. For the parenteral route of administration, the dose usually should not exceed about 1 mg, preferably about 100  $\mu$ g.

When used as vaccine agents, polynucleotides and polypeptides of the invention can be used sequentially as part of a multistep immunization process. For example, a mammal can be initially primed with a vaccine vector of the invention such as a pox virus, *e.g.*, *via* the parenteral route, and then boosted twice with the polypeptide encoded by the vaccine vector, *e.g.*, *via* the mucosal route. In another example, liposomes associated with a polypeptide or derivative of the invention can also be used for priming, with boosting being carried out mucosally using a soluble polypeptide or derivative of the invention in combination with a mucosal adjuvant (*e.g.*, LT).

A polypeptide derivative of the invention is also useful as a diagnostic reagent for detecting the presence of anti-*Chlamydia* antibodies, *e.g.*, in a blood sample. Such polypeptides are about 5 to about 80, preferably about 10 to about 50 amino acids in length and can be labeled

or unlabeled, depending upon the diagnostic method. Diagnostic methods involving such a reagent are described below.

Upon expression of a DNA molecule of the invention, a polypeptide or polypeptide derivative is produced and can be purified using known laboratory techniques. For example, the polypeptide or polypeptide derivative can be produced as a fusion protein containing a fused tail that facilitates purification. The fusion product can be used to immunize a small mammal, e.g., a mouse or a rabbit, in order to raise antibodies against the polypeptide or polypeptide derivative (monospecific antibodies). The eighth aspect of the invention thus provides a monospecific antibody that binds to a polypeptide or polypeptide derivative of the invention.

By "monospecific antibody" is meant an antibody that is capable of reacting with a unique naturally-occurring *Chlamydia* polypeptide. An antibody of the invention can be polyclonal or monoclonal. Monospecific antibodies can be recombinant, e.g., chimeric (e.g., constituted by a variable region of murine origin associated with a human constant region), humanized (a human immunoglobulin constant backbone together with hypervariable region of animal, e.g., murine, origin), and/or single chain. Both polyclonal and monospecific antibodies can also be in the form of immunoglobulin fragments, e.g., F(ab)'2 or Fab fragments. The antibodies of the invention can be of any isotype, e.g., IgG or IgA, and polyclonal antibodies can be of a single isotype or can contain a mixture of isotypes.

The antibodies of the invention, which are raised to a polypeptide or polypeptide derivative of the invention, can be produced and identified using standard immunological assays, e.g., Western blot analysis, dot blot assay, or ELISA (see, e.g., Coligan *et al.*, Current Protocols in Immunology (1994) John Wiley & Sons, Inc., New York, NY). The antibodies can be used in diagnostic methods to detect the presence of a *Chlamydia* antigen in a sample, such as a biological sample. The antibodies can also be used in affinity chromatography methods for purifying a polypeptide or polypeptide derivative of the invention. As is discussed further below, such antibodies can be used in prophylactic and therapeutic passive immunization methods.

Accordingly, a ninth aspect of the invention provides (i) a reagent for detecting the presence of *Chlamydia* in a biological sample that contains an antibody, polypeptide, or polypeptide derivative of the invention; and (ii) a diagnostic method for detecting the presence of *Chlamydia* in a biological sample, by contacting the biological sample with an antibody, a

polypeptide, or a polypeptide derivative of the invention, such that an immune complex is formed, and by detecting such complex to indicate the presence of *Chlamydia* in the sample or the organism from which the sample is derived.

Those skilled in the art will understand that the immune complex is formed between a component of the sample and the antibody, polypeptide, or polypeptide derivative, whichever is used, and that any unbound material can be removed prior to detecting the complex. As can be easily understood, a polypeptide reagent is useful for detecting the presence of anti-*Chlamydia* antibodies in a sample, *e.g.*, a blood sample, while an antibody of the invention can be used for screening a sample, such as a gastric extract or biopsy, for the presence of *Chlamydia* polypeptides.

For use in diagnostic applications, the reagent (*i.e.*, the antibody, polypeptide, or polypeptide derivative of the invention) can be in a free state or immobilized on a solid support, such as a tube, a bead, or any other conventional support used in the field. Immobilization can be achieved using direct or indirect means. Direct means include passive adsorption (non-covalent binding) or covalent binding between the support and the reagent. By "indirect means" is meant that an anti-reagent compound that interacts with a reagent is first attached to the solid support. For example, if a polypeptide reagent is used, an antibody that binds to it can serve as an anti-reagent, provided that it binds to an epitope that is not involved in the recognition of antibodies in biological samples. Indirect means can also employ a ligand-receptor system, for example, a molecule such as a vitamin can be grafted onto the polypeptide reagent and the corresponding receptor can be immobilized on the solid phase. This is illustrated by the biotin-streptavidin system. Alternatively, indirect means can be used, *e.g.*, by adding to the reagent a peptide tail, chemically or by genetic engineering, and immobilizing the grafted or fused product by passive adsorption or covalent linkage of the peptide tail.

According to a tenth aspect of the invention, there is provided a process for purifying, from a biological sample, a polypeptide or polypeptide derivative of the invention, which involves carrying out antibody-based affinity chromatography with the biological sample, wherein the antibody is a monospecific antibody of the invention.

For use in a purification process of the invention, the antibody can be polyclonal or monospecific, and preferably is of the IgG type. Purified IgGs can be prepared from an antiserum using standard methods (see, *e.g.*, Coligan *et al.*, *supra*). Conventional

chromatography supports, as well as standard methods for grafting antibodies, are disclosed in, *e.g.*, Antibodies: A Laboratory Manual, D. Lane, E. Harlow, Eds. (1988).

Briefly, a biological sample, such as an *C. pneumoniae* extract, preferably in a buffer solution, is applied to a chromatography material, preferably equilibrated with the buffer used to dilute the biological sample so that the polypeptide or polypeptide derivative of the invention (*i.e.*, the antigen) is allowed to adsorb onto the material. The chromatography material, such as a gel or a resin coupled to an antibody of the invention, can be in batch form or in a column. The unbound components are washed off and the antigen is then eluted with an appropriate elution buffer, such as a glycine buffer or a buffer containing a chaotropic agent, *e.g.*, guanidine HCl, or high salt concentration (*e.g.*, 3 M MgCl<sub>2</sub>). Eluted fractions are recovered and the presence of the antigen is detected, *e.g.*, by measuring the absorbance at 280 nm.

An antibody of the invention can be screened for therapeutic efficacy as described as follows. According to an eleventh aspect of the invention, there is provided (i) a composition of matter containing a monospecific antibody of the invention, together with a diluent or carrier; (ii) a pharmaceutical composition containing a therapeutically or prophylactically effective amount of a monospecific antibody of the invention, and (iii) a method for treating or preventing a *Chlamydia* (*e.g.*, *C. trachomatis*, *C. psittaci*, *C. pneumoniae* or *C. pecorum*) infection, by administering a therapeutic or prophylactic amount of a monospecific antibody of the invention to an individual in need. Additionally, the eleventh aspect of the invention encompasses the use of a monospecific antibody of the invention in the preparation of a medicament for treating or preventing *Chlamydia* infection.

To this end, the monospecific antibody can be polyclonal or monoclonal, preferably of the IgA isotype (predominantly). In passive immunization, the antibody can be administered to a mucosal surface of a mammal, *e.g.*, the gastric mucosa, *e.g.*, orally or intragastrically, advantageously, in the presence of a bicarbonate buffer. Alternatively, systemic administration, not requiring a bicarbonate buffer, can be carried out. A monospecific antibody of the invention can be administered as a single active component or as a mixture with at least one monospecific antibody specific for a different *Chlamydia* polypeptide. The amount of antibody and the particular regimen used can be readily determined by one skilled in the art. For example, daily administration of about 100 to 1,000 mg of antibodies over one week, or three doses per day of

about 100 to 1,000 mg of antibodies over two or three days, can be an effective regimens for most purposes.

Therapeutic or prophylactic efficacy can be evaluated using standard methods in the art, *e.g.*, by measuring induction of a mucosal immune response or induction of protective and/or therapeutic immunity, using, *e.g.*, the *C. pneumoniae* mouse model. Those skilled in the art will recognize that the *C. pneumoniae* strain of the model can be replaced with another *Chlamydia* strain. For example, the efficacy of DNA molecules and polypeptides from *C. pneumoniae* is preferably evaluated in a mouse model using an *C. pneumoniae* strain. Protection can be determined by comparing the degree of *Chlamydia* infection to that of a control group. Protection is shown when infection is reduced by comparison to the control group. Such an evaluation can be made for polynucleotides, vaccine vectors, polypeptides and derivatives thereof, as well as antibodies of the invention.

Adjuvants useful in any of the vaccine compositions described above are as follows.

Adjuvants for parenteral administration include aluminum compounds, such as aluminum hydroxide, aluminum phosphate, and aluminum hydroxy phosphate. The antigen can be precipitated with, or adsorbed onto, the aluminum compound according to standard protocols. Other adjuvants, such as RIBI (ImmunoChem, Hamilton, MT), can be used in parenteral administration.

Adjuvants for mucosal administration include bacterial toxins, *e.g.*, the cholera toxin (CT), the *E. coli* heat-labile toxin (LT), the *Clostridium difficile* toxin A and the *pertussis* toxin (PT), or combinations, subunits, toxoids, or mutants thereof. For example, a purified preparation of native cholera toxin subunit B (CTB) can be of use. Fragments, homologs, derivatives, and fusions to any of these toxins are also suitable, provided that they retain adjuvant activity. Preferably, a mutant having reduced toxicity is used. Suitable mutants are described, *e.g.*, in WO 95/17211 (Arg-7-Lys CT mutant), WO 96/6627 (Arg-192-Gly LT mutant), and WO 95/34323 (Arg-9-Lys and Glu-129-Gly PT mutant). Additional LT mutants that can be used in the methods and compositions of the invention include, *e.g.*, Ser-63-Lys, Ala-69-Gly, Glu-110-Asp, and Glu-112-Asp mutants. Other adjuvants, such as a bacterial monophosphoryl lipid A (MPLA) of, *e.g.*, *E. coli*, *Salmonella minnesota*, *Salmonella typhimurium*, or *Shigella flexneri*; saponins, or polylactide glycolide (PLGA) microspheres, can also be used in mucosal administration.

Adjuvants useful for both mucosal and parenteral administrations include polyphosphazene (WO 95/2415), DC-chol (3 b-(N-(N',N'-dimethyl aminomethane)-carbamoyl) cholesterol; U.S. Patent No. 5,283,185 and WO 96/14831) and QS-21 (WO 88/9336).

Any pharmaceutical composition of the invention, containing a polynucleotide, a polypeptide, a polypeptide derivative, or an antibody of the invention, can be manufactured in a conventional manner. In particular, it can be formulated with a pharmaceutically acceptable diluent or carrier, *e.g.*, water or a saline solution such as phosphate buffer saline. In general, a diluent or carrier can be selected on the basis of the mode and route of administration, and standard pharmaceutical practice. Suitable pharmaceutical carriers or diluents, as well as pharmaceutical necessities for their use in pharmaceutical formulations, are described in *Remington's Pharmaceutical Sciences*, a standard reference text in this field and in the USP/NF.

The invention also includes methods in which *Chlamydia* infection, are treated by oral administration of a *Chlamydia* polypeptide of the invention and a mucosal adjuvant, in combination with an antibiotic, an antacid, sucralfate, or a combination thereof. Examples of such compounds that can be administered with the vaccine antigen and the adjuvant are antibiotics, including, *e.g.*, macrolides, tetracyclines, and derivatives thereof (specific examples of antibiotics that can be used include azithromycin or doxycyclin or immunomodulators such as cytokines or steroids. In addition, compounds containing more than one of the above-listed components coupled together, can be used. The invention also includes compositions for carrying out these methods, *i.e.*, compositions containing a *Chlamydia* antigen (or antigens) of the invention, an adjuvant, and one or more of the above-listed compounds, in a pharmaceutically acceptable carrier or diluent.

Amounts of the above-listed compounds used in the methods and compositions of the invention can readily be determined by one skilled in the art. In addition, one skilled in the art can readily design treatment/immunization schedules. For example, the non-vaccine components can be administered on days 1-14, and the vaccine antigen + adjuvant can be administered on days 7, 14, 21, and 28.

# REFERENCES

1. Grayston et al. (1995) Journal of Infectious Diseases 168:1231
2. Campos et al. (1995) Investigation of Ophthalmology and Visual Science 36:1477
3. Grayston et al (1990) Journal of Infectious Diseases 161:618
- 5 4. Marrie (1993) Clinical Infectious Diseases. 18:501
5. Wang et al (1986) Chlamydial infections. Cambridge University Press, Cambridge. p. 329
6. Saikku et al.(1988) Lancet;ii:983
7. Thom et al. (1992) JAMA 268:68
8. Linnanmaki et al. (1993), Circulation 87:1030
- 10 9. Saikku et al. (1992)Annals Internal Medicine 116:273
10. Melnick et al(1993) American Journal of Medicine 95:499
- 11 Shor et al. (1992) South African. Medical Journal 82:158
12. Kuo et al. (1993) Journal of Infectious Diseases 167:841
13. Kuo et al. (1993) Arteriosclerosis and Thrombosis 13:1500
- 15 14. Campbell et al (1995) Journal of Infectious Diseases 172:585
15. Chiu et al. Circulation, 1997 (In Press).
16. Ramirez et al (1996) Annals of Internal Medicine 125:979
17. Jackson et al. Abst. K121, p272, 36th ICAAC, 15-18 Sept. 1996, New Orleans.
18. Fong et al (1997) Journal of Clinical Microbiology 35:48
- 20 19. Hahn DL, et al. Evidence for Chlamydia pneumoniae infection in steroid-dependent asthma.  
Ann Allergy Asthma Immunol. 1998 Jan; 80(1): 45-49.
20. Hahn DL, et al. Association of Chlamydia pneumoniae IgA antibodies with recently  
symptomatic asthma. Epidemiol Infect. 1996 Dec; 117(3): 513-517.
21. Bjornsson E, et al. Serology of chlamydia in relation to asthma and bronchial  
hyperresponsiveness. Scand J Infect Dis. 1996; 28(1): 63-69.
- 25 22. Hahn DL. Treatment of *Chlamydia pneumoniae* infection in adult asthma: a before-after trial.  
J Fam Pract. 1995 Oct; 41(4): 345-351.
23. Allegra L, et al. Acute exacerbations of asthma in adults: role of Chlamydia pneumoniae  
infection. Eur Respir J. 1994 Dec; 7(12): 2165-2168.
- 30 24. Hahn DL, et al. Association of Chlamydia pneumoniae (strain TWAR) infection with  
wheezing, asthmatic bronchitis, and adult-onset asthma. JAMA. 1991 Jul 10; 266(2): 225-230.



25. Pal *et al.* (1996) *Infection and Immunity* 64:5341
26. Jones *et al.* (1995) *Vaccine* 13:715
27. Igiertsemes *et al.* (1993) *Immunology* 5:317
28. Igiertseme *et al.* (1993) *Regional Immunology* 5:317
- 5 29. Magee *et al.* (1993) *Regional Immunology* 5: 305
30. Landers *et al.* (1991) *Infection & Immunity* 59:3774
31. Magee *et al.* (1995) *Infection & Immunity* 63:516
32. Cotter *et al.* (1995) *Infection and Immunity* 63:4704
33. Campbell *et al.* (1990) *Infection and Immunity* 58:93
- 10 34. McCafferty *et al.* (1995) *Infection and Immunity* 63:2387-9.
35. Knudsen *et al.* (1996) Third Meeting of the European Society for Chlamydia Research, Vienna
36. Wiedmann-Al-Ahmad M, *et al.* Reactions of polyclonal and neutralizing anti-p54 monoclonal antibodies with an isolated, species-specific 54-kilodalton protein of *Chlamydia pneumoniae*. *Clin Diagn Lab Immunol.* 1997 Nov; 4(6): 700-704
- 15 37. Hughes *et al.*, 1992. *Infect. Immun.* 60(9):3497,
38. Dion *et al.*, 1990. *Virology* 179:474-477,
39. Snijders *et al.*, 1991. *J. Gen. Virol.* 72:557-565,
40. Langeveld *et al.*, *Vaccine* 12(15):1473-1480, 1994
41. Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley & Sons Inc., 1994
- 20 42. Kunkel *et al.* *Proc. Natl. Acad. Sci. USA* (1985) 82:448
43. Silhavy *et al.* *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory Press, 1984
44. Davis *et al.* *A Manual for Genetic Engineering: Advanced Bacterial Genetics*, Cold Spring Harbor Laboratory Press, 1980)
45. Casey & Davidson, *Nucl. Acid Res.* (1977) 4:1539
- 25 46. Cagnon *et al.*, *Protein Engineering* (1991) 4(7):843
47. Takase *et al.*, *J. Bact.* (1987) 169:5692

gataaaaaatt cttgacagct gttttgtcat ctttaacttg atttacttat tttgtttcta															60
tattgatgcg aatagtcttc taaaaaaciaa aagcattacc															115
Met Lys Thr Ser Ile															
1 5															
cct tgg gtt tta gtt tcc tcc gtg tta gct ttc tca tgt cac cta cag															163
Pro Trp Val Leu Val Ser Ser Val Leu Ala Phe Ser Cys His Leu Gln															
10 15 20															
tca cta gct aac gag gaa ctt tta tca cct gat gat agc ttt aat gga															211
Ser Leu Ala Asn Glu Glu Leu Leu Ser Pro Asp Asp Ser Phe Asn Gly															
25 30 35															
aat atc gat tca gga acg ttt act cca aaa act tca gcc aca aca tat															259
Asn Ile Asp Ser Gly Thr Phe Thr Pro Lys Thr Ser Ala Thr Thr Tyr															
40 45 50															
tct cta aca gga gat gtc ttc ttt tac gag cct gga aaa ggc act ccc															307
Ser Leu Thr Gly Asp Val Phe Phe Tyr Glu Pro Gly Lys Gly Thr Pro															
55 60 65															
tta tct gac agt tgt ttt aag caa acc acg gac aat ctt acc ttc ttg															355
Leu Ser Asp Ser Cys Phe Lys Gln Thr Thr Asp Asn Leu Thr Phe Leu															
70 75 80 85															
ggg aac ggt cat agc tta acg ttt ggc ttt ata gat gct ggc act cat															403
Gly Asn Gly His Ser Leu Thr Phe Gly Phe Ile Asp Ala Gly Thr His															
90 95 100															
gca ggt gct gct gca tct aca aca gca aat aag aat ctt acc ttc tca															451
Ala Gly Ala Ala Ala Ser Thr Thr Ala Asn Lys Asn Leu Thr Phe Ser															
105 110 115															
ggg ttt tcc tta ctg agt ttt gat tcc tct cct agc aca acg gtt act															499
Gly Phe Ser Leu Leu Ser Phe Asp Ser Ser Pro Ser Thr Thr Val Thr															
120 125 130															
aca ggt cag gga acg ctt tcc tca gca gga ggc gta aat tta gaa aat															547
Thr Gly Gln Gly Thr Leu Ser Ser Ala Gly Gly Val Asn Leu Glu Asn															
135 140 145															
att cgt aaa ctt gta gtt gct ggg aat ttt tct act gca gat ggt gga															595
Ile Arg Lys Leu Val Val Ala Gly Asn Phe Ser Thr Ala Asp Gly Gly															
150 155 160 165															
gct atc aaa gga gcg tct ttc ctt tta act ggc act tct gga gat gct															643
Ala Ile Lys Gly Ala Ser Phe Leu Leu Thr Gly Thr Ser Gly Asp Ala															
170 175 180															
ctt ttt agt aac aac tct tca tca aca aag gga gga gca att gct act															691
Leu Phe Ser Asn Asn Ser Ser Ser Thr Lys Gly Gly Ala Ile Ala Thr															
185 190 195															
aca gca ggc gct cgc ata gca aat aac aca ggt tat gtt aga ttc cta															739
Thr Ala Gly Ala Arg Ile Ala Asn Asn Thr Gly Tyr Val Arg Phe Leu															
200 205 210															
tct aac ata gcg tct acg tca gga ggc gct atc gat gat gaa ggc acg															787
Ser Asn Ile Ala Ser Thr Ser Gly Gly Ala Ile Asp Asp Glu Gly Thr															
215 220 225															

tcg ata cta tcg aac aac aaa ttt cta tat ttt gaa ggg aat gca gcg 835  
 Ser Ile Leu Ser Asn Asn Lys Phe Leu Tyr Phe Glu Gly Asn Ala Ala  
 230 235 240 245  
 aaa act act ggc ggt gcg atc tgc aac acc aag gcg agt gga tct cct 883  
 Lys Thr Thr Gly Gly Ala Ile Cys Asn Thr Lys Ala Ser Gly Ser Pro  
 250 255 260  
 gaa ctg ata atc tct aac aat aag act ctg atc ttt gct tca aac gta 931  
 Glu Leu Ile Ile Ser Asn Asn Lys Thr Leu Ile Phe Ala Ser Asn Val  
 265 270 275  
 gca gaa aca agc ggt ggc gcc atc cat gct aaa aag cta gcc ctt tcc 979  
 Ala Glu Thr Ser Gly Gly Ala Ile His Ala Lys Lys Leu Ala Leu Ser  
 280 285 290  
 tct gga ggc ttt aca gag ttt cta cga aat aat gtc tca tca gca act 1027  
 Ser Gly Gly Phe Thr Glu Phe Leu Arg Asn Asn Val Ser Ser Ala Thr  
 295 300 305  
 cct aag ggg ggt gct atc agc atc gat gcc tca gga gag ctc agt ctt 1075  
 Pro Lys Gly Gly Ala Ile Ser Ile Asp Ala Ser Gly Glu Leu Ser Leu  
 310 315 320 325  
 tct gca gag aca gga aac att acc ttt gta aga aat acc ctt aca aca 1123  
 Ser Ala Glu Thr Thr Gly Asn Ile Thr Phe Val Arg Asn Thr Leu Thr Thr  
 330 335 340  
 acc gga agt acc gat act cct aaa cgt aat gcg atc aac ata gga agt 1171  
 Thr Gly Ser Thr Asp Thr Pro Lys Arg Asn Ala Ile Asn Ile Gly Ser  
 345 350 355  
 aac ggg aaa ttc acg gaa tta cgg gct gct aaa aat cat aca att ttc 1219  
 Asn Gly Lys Phe Thr Glu Leu Arg Ala Ala Lys Asn His Thr Ile Phe  
 360 365 370  
 ttc tat gat ccc atc act tca gaa gga acc tca tca gac gta ttg aag 1267  
 Phe Tyr Asp Pro Ile Thr Ser Glu Gly Thr Ser Ser Asp Val Leu Lys  
 375 380 385  
 ata aat aac ggc tct gcg gga gct ctc aat cca tat caa gga acg att 1315  
 Ile Asn Asn Gly Ser Ala Gly Ala Leu Asn Pro Tyr Gln Gly Thr Ile  
 390 395 400 405  
 cta ttt tct gga gaa acc cta aca gca gat gaa ctt aaa gtt gct gac 1363  
 Leu Phe Ser Gly Glu Thr Leu Thr Ala Asp Glu Leu Lys Val Ala Asp  
 410 415 420  
 aat tta aaa tct tca ttc acg cag cca gtc tcc cta tcc gga gga aag 1411  
 Asn Leu Lys Ser Ser Phe Thr Gln Pro Val Ser Leu Ser Gly Gly Lys  
 425 430 435  
 tta ttg cta caa aag gga gtc act tta gag agc acg agc ttc tct caa 1459  
 Leu Leu Leu Gln Lys Gly Val Thr Leu Glu Ser Thr Ser Phe Ser Gln  
 440 445 450  
 gag gcc ggt tct ctc ctc ggc atg gat tca gga acg aca tta tca act 1507  
 Glu Ala Gly Ser Leu Leu Gly Met Asp Ser Gly Thr Thr Leu Ser Thr  
 455 460 465  
 aca gct ggg agt att aca atc acg aac cta gga atc aat gtt gac tcc 1555  
 Thr Ala Gly Ser Ile Thr Ile Thr Asn Leu Gly Ile Asn Val Asp Ser  
 470 475 480 485

tta ggt ctt aag cag ccc gtc agc cta aca gca aaa ggt gct tca aat 1603  
 Leu Gly Leu Lys Gln Pro Val Ser Leu Thr Ala Lys Gly Ala Ser Asn  
 490 495 500

aaa gtg atc gta tct ggg aag ctc aac ctg att gat att gaa ggg aac 1651  
 Lys Val Ile Val Ser Gly Lys Leu Asn Leu Ile Asp Ile Glu Gly Asn  
 505 510 515

att tat gaa agt cat atg ttc agc cat gac cag ctc ttc tct cta tta 1699  
 Ile Tyr Glu Ser His Met Phe Ser His Asp Gln Leu Phe Ser Leu Leu  
 520 525 530

aaa atc acg gtt gat gct gat gtt gat act aac gtt gac atc agc agc 1747  
 Lys Ile Thr Val Asp Ala Asp Val Asp Thr Asn Val Asp Ile Ser Ser  
 535 540 545

ctt atc cct gtt cct gct gag gat cct aat tca gaa tac gga ttc caa 1795  
 Leu Ile Pro Val Pro Ala Glu Asp Pro Asn Ser Glu Tyr Gly Phe Gln  
 550 555 560 565

gga caa tgg aat gtt aat tgg act acg gat aca gct aca aat aca aaa 1843  
 Gly Gln Trp Asn Val Asn Trp Thr Thr Asp Thr Ala Thr Asn Thr Lys  
 570 575 580

gag gcc acg gca act tgg acc aaa aca gga ttt gtt ccc agc ccc gaa 1891  
 Glu Ala Thr Ala Thr Trp Thr Lys Thr Gly Phe Val Pro Ser Pro Glu  
 585 590 595

aga aaa tct gcg tta gta tgc aat acc cta tgg gga gtc ttt act gac 1939  
 Arg Lys Ser Ala Leu Val Cys Asn Thr Leu Trp Gly Val Phe Thr Asp  
 600 605 610

att cgc tct ctg caa cag ctt gta gag atc ggc gca act ggt atg gaa 1987  
 Ile Arg Ser Leu Gln Gln Leu Val Glu Ile Gly Ala Thr Gly Met Glu  
 615 620 625

cac aaa caa ggt ttc tgg gtt tcc tcc atg acg aac ttc ctg cat aag 2035  
 His Lys Gln Gly Phe Thr Val Ser Ser Met Thr Asn Phe Leu His Lys  
 630 635 640 645

act gga gat gaa aat cgc aaa ggc ttc cgt cat acc tct gga ggc tac 2083  
 Thr Gly Asp Glu Asn Arg Lys Gly Phe Arg His Thr Ser Gly Gly Tyr  
 650 655 660

gtc atc ggt gga agt gct cac act cct aaa gac gac cta ttt acc ttt 2131  
 Val Ile Gly Gly Ser Ala His Thr Pro Lys Asp Asp Leu Phe Thr Phe  
 665 670 675

gcg ttc tgc cat ctc ttt gct aga gac aaa gat tgt ttt atc gct cac 2179  
 Ala Phe Cys His Leu Phe Ala Arg Asp Lys Asp Cys Phe Ile Ala His  
 680 685 690

aac aac tct aga acc tac ggt gga act tta ttc ttc aag cac tct cat 2227  
 Asn Asn Ser Arg Thr Tyr Gly Gly Thr Leu Phe Phe Lys His Ser His  
 695 700 705

acc cta caa ccc caa aac tat ttg aga tta gga aga gca aag ttt tct 2275  
 Thr Leu Gln Pro Gln Asn Tyr Leu Arg Leu Gly Arg Ala Lys Phe Ser  
 710 715 720 725

gaa tca gct ata gaa aaa ttc cct agg gaa att ccc cta gcc ttg gat 2323  
 Glu Ser Ala Ile Glu Lys Phe Pro Arg Glu Ile Pro Leu Ala Leu Asp  
 730 735 740

gtc caa gtt tcg ttc agc cat tca gac aac cgt atg gaa acg cac tat 2371  
 Val Gln Val Ser Phe Ser His Ser Asp Asn Arg Met Glu Thr His Tyr  
 745 750 755

acc tca ttg cca gaa tcc gaa ggt tct tgg agc aac gag tgt ata gct 2419  
 Thr Ser Leu Pro Glu Ser Glu Gly Ser Trp Ser Asn Glu Cys Ile Ala  
 760 765 770

ggt ggt atc ggc cta gac ctt cct ttt gtt ctt tcc aac cca cat cct 2467  
 Gly Gly Ile Gly Leu Asp Leu Pro Phe Val Leu Ser Asn Pro His Pro  
 775 780 785

ctt ttc aag acc ttc att cca cag atg aaa gtc gaa atg gtt tat gta 2515  
 Leu Phe Lys Thr Phe Ile Pro Gln Met Lys Val Glu Met Val Tyr Val  
 790 795 800 805

tca caa aat agc ttc ttc gaa agc tct agt gat ggc cgt ggt ttt agt 2563  
 Ser Gln Asn Ser Phe Phe Glu Ser Ser Ser Asp Gly Arg Gly Phe Ser  
 810 815 820

att gga agg ctg ctt aac ctc tcg att cct gtg ggt gcg aaa ttc gtg 2611  
 Ile Gly Arg Leu Leu Asn Leu Ser Ile Pro Val Gly Ala Lys Phe Val  
 825 830 835

cag ggg gat atc gga gat tcc tac acc tat gat ctc tca gga ttc ttt 2659  
 Gln Gly Asp Ile Gly Asp Ser Tyr Thr Tyr Asp Leu Ser Gly Phe Phe  
 840 845 850

gtt tcc gat gtc tat cgt aac aat ccc caa tct aca gcg act ctt gtg 2707  
 Val Ser Asp Val Tyr Arg Asn Asn Pro Gln Ser Thr Ala Thr Leu Val  
 855 860 865

atg agc cca gac tct tgg aaa att cgc ggt ggc aat ctt tca aga cag 2755  
 Met Ser Pro Asp Ser Trp Lys Ile Arg Gly Gly Asn Leu Ser Arg Gln  
 870 875 880 885

gca ttt tta ctg agg ggt agc aac aac tac gtc tac aac tcc aat tgt 2803  
 Ala Phe Leu Leu Arg Gly Ser Asn Asn Tyr Val Tyr Asn Ser Asn Cys  
 890 895 900

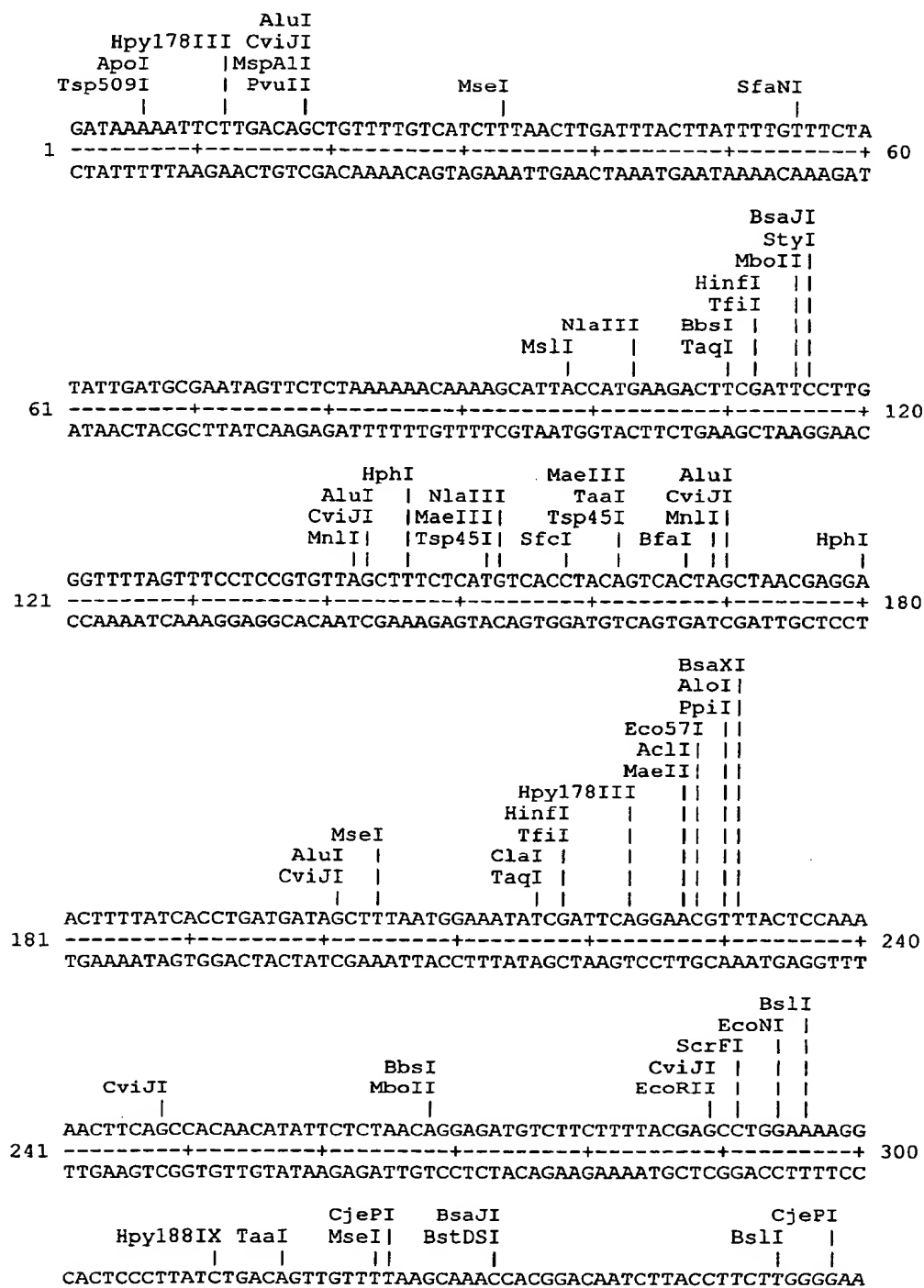
gag ctc ttc gga cat tac gct atg gaa ctc cgt gga tct tca agg aac 2851  
 Glu Leu Phe Gly His Tyr Ala Met Glu Leu Arg Gly Ser Ser Arg Asn  
 905 910 915

tac aat gta gat gtt ggt acc aaa ctc cga ttc tagattgcta aaactcccta 2904  
 Tyr Asn Val Asp Val Gly Thr Lys Leu Arg Phe  
 920 925

gttcttctag ggagttttct catactttta gggaaatatt tgctat 2950

Figure 2

## Restriction enzyme analysis of CPN100634



301 -----+-----+-----+-----+-----+ 360  
 GTGAGGGAATAGACTGTCAACAAATTCGTTTGGTGCCTGTTAGAATGGAAGAACCCCTT

CviRI  
 Fnu4HI |  
 TseI | |  
 Fnu4HI | | |  
 AlwNI | | |  
 BstAPI | | |  
 MwoI | | |  
 TseI | | |  
 AarI | | | |  
 MslI | | | |  
 CviRI | | | |  
 NlaIII | | | |  
 BbvI | | | |  
 BbvI | | | |  
 BspMI | | | |  
 Cac8I | | | |

AclI  
 MaeII  
 MseI |  
 AluI | | CviJI  
 TaaI CviJI | | SfaNI |

361 -----+-----+-----+-----+-----+ 420  
 CGGTCATAGCTTAACGTTTGGCTTTATAGATGCTGGCACTCATGCAGGTGCTGCTGCATC  
 GCCAGTATCGAATTGCAAACCGAAATATCTACGACCGTGAGTACGTCCACGACGACGTAG

BseMII  
 SfaNI      HinfI      BslI      DdeI      HinfI  
 TfiI      DdeI      BseMII      TfiI

421 -----+-----+-----+-----+-----+ 480  
 TACAACAGCAAATAAGAATCTTACCTTCTCAGGGTTTTCCTTACTGAGTTTGTATTCCTC  
 ATGTTGTCGTTTATTCTTAGAATGGAAGAGTCCCAAAGGAATGACTCAAAACTAAGGAG

MaeIII      MnlI      ApoI  
 BfaI      TaaI      SfcI      XmnI      DdeI      MnlI      BseMII  
 BbvCI |      Tsp509I  
 Bpu10I |

481 -----+-----+-----+-----+-----+ 540  
 TCCTAGCACAACGGTTACTACAGGTCAGGGAACGCTTTCCTCAGCAGGAGGCGTAAATTT  
 AGGATCGTGTGCAATGATGTCCAGTCCCTTGCAGAAAGGAGTCGTCTCCGCATTTAA

SspI      ApoI      BccI      HgaI  
 XmnI |      CjeI      PstI |      AluI |  
 Tsp509I      CviRI |      SfcI |      CviJI |

541 -----+-----+-----+-----+-----+ 600  
 AGAAAAATATTCGTAAACTTGTAGTTGCTGGGAATTTTCTACTGCAGATGGTGGAGCTAT  
 TCTTTTATAAGCATTTGAACATCAACGACCCTTAAAAAGATGACGTCTACCACCTCGATA

Hpy178III      BpmI  
 CjeI      MseI      BsrI      MboII |  
 SfaNI |      MaeIII |

601 -----+-----+-----+-----+-----+ 660  
 CAAAGGAGCGTCTTTCCTTTTAACTGGCACTTCTGGAGATGCTCTTTTATAGTAACAACCTC  
 GTTTCCTCGCAGAAAGGAAAATTGACCGTGAAGACCTCTACGAGAAAAATCATTGTTGAG

Cac8I  
 HaeII |  
 HhaI | |  
 MnlI      MunI      BseRI |      Cac8I      MwoI  
 EarI |      Tsp509I      SfcI |

TTCATCAACAAAGGGAGGAGCAATTGCTACTACAGCAGGCGCTCGCATAGCAAATAACAC

38



BsiHKAI |  
 Bsp1286I |  
 SacI |  
 Hpy178III DdeI |  
 Bsu36I | AluI | CviRI  
 ClaI DdeI | CviJI | SfcI |  
 Bsu36I DdeI BslI SfaNI TaqI SfaNI | MnlI | BsmAI |  
 AGCAACTCCTAAGGGGGTGTATCAGCATCGATGCCTCAGGAGAGCTCAGTCTTTCTGC  
 1021 -----+----- 1080  
 TCGTTGAGGATTCCCCCAGATAGTCGTAGCTACGGAGTCCTCTCGAGTCAGAAAGACG  
 BseMII MspI BsaXI  
 PstI BsaWI | RsaI |  
 AGAGACAGGAAACATTACCTTTGTAAGAAATACCCTTACAACAACCGAAGTACCGATAC  
 1081 -----+----- 1140  
 TCTCTGTCCTTTGTAATGGAAACATTCTTTATGGGAATGTTGTTGGCCTTCATGGCTATG  
 Tsp509I  
 Sth132I | Fnu4HI  
 ApoI | CviJI |  
 DpnI MaeIII Tsp509I | TseI |  
 MaeII Sau3AI | Sth132I | BscGI | BbvI | BscGI |  
 TCCTAAACGTAATGCGATCAACATAGGAAGTAACGGGAAATTCACGGAATTACGGGCTGC  
 1141 -----+----- 1200  
 AGGATTTCATTACGCTAGTTGTATCCTTCATTGCCCTTTAAGTGCCTTAATGCCCGACG  
 DpnI MnlI  
 Sau3AI | MaeII |  
 Tsp509I Eco57I | Hpy188IX Hpy188IX |  
 MboII | AlwI | BccI | NlaIV |  
 TAAAAATCATACAATTTTCTTCTATGATCCCATCACTTCAGAAGGAACCTCATCAGACGT  
 1201 -----+----- 1260  
 ATTTTATGATGTAAAAGAAGATACTAGGGTAGTGAAGTCTTCCTTGGAGTAGTCTGCA  
 BanII  
 BceFI  
 BsiHKAI  
 CviJI Bsp1286I  
 FauI | SacI  
 MboII | AluI |  
 Sth132I | AciI CviJI | BslI HinfI  
 TfiI  
 ATTGAAGATAAATAACGGCTCTGCGGGAGCTCTCAATCCATATCAAGGAACGATTCTATT  
 1261 -----+----- 1320  
 TAACTTCTATTTATTGCCGAGACGCCCTCGAGAGTTAGGTATAGTTCCTTGCTAAGATAA  
 DraI  
 MseI |  
 MboII |  
 Hpy178III BpmI MseI Tsp509I |  
 TTCTGGAGAAACCCTAACAGCAGATGAACCTAAAGTTGCTGACAATTTAAATCTTCATT  
 1321 -----+----- 1380  
 AAGACCTCTTTGGGATTGTCGTCTACTTGAATTTCAACGACTGTAAATTTTAGAAGTAA  
 Hpy178III  
 MspI |  
 BsaWI |  
 BsrI BspEI |  
 CviJI | MnlI |  
 MaeIII

Fnu4HI || BbvI || || Tsp45I Bce83I  
 TseI || BsmAI || || HinfI PleI  
 || || || || ||  
 CACGCAGCCAGTCTCCCTATCCGGAGGAAAGTTATTGCTACAAAAGGGAGTCACTTTAGA  
 1381 -----+-----+-----+-----+-----+ 1440  
 GTGCGTCGGTCAGAGGGATAGGCCTCCTTTCAATAACGATGTTTTCCCTCAGTGAAATCT

MspI  
 BseRI  
 BsrFI  
 CviJI  
 HaeIII  
 Hpy178III  
 SmlI  
 MnlI  
 HinfI  
 AluI  
 CviJI  
 BsiHKA  
 Bsp1286I  
 BssSI  
 Hpy178III  
 HinfI  
 MnlI  
 TfiI  
 BsaJI  
 NlaIII  
 BplI  
 BslI  
 CjeI  
 GAGCAGAGCTTCTCTCAAGAGGCCGGTTCTCTCCTCGGCATGGATTGAGGAACGACATT  
 1441 -----+-----+-----+-----+-----+ 1500  
 CTCGTGCTCGAAGAGAGTTCTCCGCCAAGAGAGGAGCCGTACCTAAGTCCTTGCTGTAA

HinfI  
 BfaI  
 Bsu36I  
 AluI  
 CviJI  
 MspAII  
 PvuII  
 SfcI  
 Hpy178III  
 StyI  
 TfiI  
 PleI  
 HincII  
 DdeI  
 ATCAACTACAGCTGGGAGTATTACAATCACGAACCTAGGAATCAATGTTGACTCCTTAGG  
 1501 -----+-----+-----+-----+-----+ 1560  
 TAGTTGATGTCGACCCTCATAATGTTAGTGCTTGGATCCTTAGTTACAAGTGAAGGATCC

BbvI  
 CviJI  
 Sth132I  
 BscGI  
 CviJI  
 Fnu4HI  
 MseI  
 AflIII  
 SmlI  
 TseI  
 BstAPI  
 MwoI  
 CjePI  
 Sau3AI  
 DpnI  
 TCTTAAGCAGCCCGTCAGCCTAACAGCAAAGGTGCTTCAAATAAAGTGATCGTATCTGG  
 1561 -----+-----+-----+-----+-----+ 1620  
 AGAATTCGTCGGGCAGTCGGATTGTCGTTTTCCACGAAGTTTATTTCACTAGCATAGACC

NlaIII  
 MboII  
 AluI  
 CviJI  
 CjePI  
 NdeI  
 CviJI  
 CjePI  
 GAAAGCTCAACCTGATTGATATTGAAGGGAACATTTATGAAAGTCATATGTTTCAGCCATGA  
 1621 -----+-----+-----+-----+-----+ 1680  
 CTTGAGTTGGACTAACTATAACTTCCCTTGTAATACTTTTCAGTATACAAGTCGGTACT

MseI  
 AceIII  
 TaaI  
 HincII  
 AluI  
 EarI  
 CjePI  
 AclI  
 CviJI  
 SapI  
 SfaNI  
 BaeI  
 MaeII

1681 CCAGCTCTTCTCTCTATTAAAAATCACGGTTGATGCTGATGTTGATACTAACGTTGACAT  
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1740  
 GGTCGAGAAGAGAGATAATTTTGTAGTGCCAACTACGACTACAACATGATTGCAACTGTA

Hpy188IX  
 AlwI |  
 Tsp509I | |  
 DpnI | |  
 NlaIV | |  
 BamHI | |  
 BstYI | |  
 Sau3AI | |  
 BbvCI | |  
 Bpu10I | |  
 BaeI | |  
 CviJI | | MnlI | |  
 Fnu4HI | | BbvI | | DdeI | |  
 TseI | | BseMII | | AlwI | |  
 BsaJI | | StyI | |  
 HinfI | |  
 TfiI | |

1741 CAGCAGCCTTATCCCTGTTCTGCTGAGGATCCTAATTCAGAATACGGATTCCAAGGACA  
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1800  
 GTCGTCGGAATAGGGACAAGGACGACTCCTAGGATTAAGTCTTATGCCTAAGGTTCTCTGT

BsaJI  
 BstDSI  
 CviJI |  
 HaeI |  
 Tsp509I | |  
 MseI | BciVI |  
 AluI |  
 CviJI | MnlI |  
 HaeIII |  
 AvaII  
 Sau96I

1801 ATGGAATGTTAATTGGACTACGGATACAGCTACAAATACAAAAGAGGCCACGGCAACTTG  
 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1860  
 TACCTTACAATTAACCTGATGCCTATGTCGATGTTTATGTTTCTCCGGTGCCGTTGAAC

Bcefi |  
 CviJI | Sth132I |  
 CviRI |  
 GACCAAAACAGGATTTGTTCCAGCCCCGAAAGAAAATCTGCGTTAGTATGCAATACCTT  
 1861 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1920  
 CTGGTTTTGTCTAAACAAGGGTCGGGGCTTTCTTTTAGACGCAATCATACGTTATGGGA

HinfI | PleI | Hin4I |  
 CviRI | CviJI | Sau3AI |  
 DpnI |  
 HhaI |  
 ATGGGGAGTCTTTACTGACATTGCTCTCTGCAACAGCTTGTAGAGATCGGCGCAACTGG  
 1921 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 1980  
 TACCCCTCAGAAATGACTGTAAGCGAGAGACGTTGTCGAACATCTCTAGCCGCGTTGACC

BcgI | MnlI |  
 BsrI | Tth111III | NlaIII |  
 CviRI |  
 TATGGAACACAAACAAGGTTTCTGGGTTTCTCCTCCATGACGAACCTCCTGCATAAGACTGG  
 1981 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2040  
 ATACCTTGTGTTTGTTCCAAAGACCCAAAGGAGGTACTGCTTGAAGGACGTATTCTGACC

MaeII  
 BsaXI |  
 Hin4I |  
 BsaBI |  
 BcgI | BpmI |  
 BsrI | CviJI | MnlI |  
 Hpy178III |  
 CviJI | BpmI |  
 AGATGAAAATCGCAAAGGCTTCCGTCATACCTCTGGAGGCTACGTCATCGGTGGAAGTGC  
 2041 -----+-----+-----+-----+-----+-----+-----+-----+-----+-----+ 2100  
 TCTACTTTTAGCGTTTCCGAAGGCAGTATGGAGACCTCCGATGCAGTAGCCACCTTCACG

BsiHKAI

BfaI

42

43

```

          Sau3AI |
        BsaJI  | |
        BstDSI | |
        MboII  | |   AlwI
          ||   | |
2821  CGCTATGGAACTCCGTGGATCTTCAAGGAACTACAATGTAGATGTTGGTACCAAACCTCCG
-----+-----+-----+-----+-----+ 2880
        GCGATACCTTGAGGCACCTAGAAGTTCCTTGATGTTACATCTACAACCATGGTTTGAGGC

Hpy178III
  BfaI |
XbaI | |
  |||
2881  ATTCTAGATTGCTAAAACTCCCTAGTTCTTCTAGGGAGTTTTCTCATACTTTTAGGGAAA
-----+-----+-----+-----+ 2940
        TAAGATCTAACGATTTTGAGGGATCAAGAAGATCCCTCAAAGAGTATGAAAATCCCTTT

SspI
  |
  TATTTGCTAT
2941 -----+ 2950
        ATAAACGATA

```

861021 220109